

**ANAPHORIC RESOLUTION OF ZERO PRONOUNS IN  
CHINESE IN TRANSLATION AND READING  
COMPREHENSION**

**VALERIE PELLATT**

**Ph. D. Thesis**

**UNIVERSITY OF EDINBURGH**

**2002**



## DECLARATION

I hereby declare that this thesis is entirely of my own composition, that it describes my own research, and that it has not been submitted to this or any other university for a degree, in this or any other form.

Valerie Pellatt



## ABSTRACT OF THE THESIS

The primary aim of the thesis is to investigate some of the processes of reading Chinese text by means of comparing and analysing approximately 100 parallel translations of four texts from Chinese to English. The translations are answers to A Level examination questions. The focus of the investigation is interpretation of the zero pronoun, a common phenomenon in Chinese, which often requires explicitation when translated into English. The secondary aim is to show how translation gives evidence of comprehension, as shown by the variation in interpretation of zero pronouns. The thesis reviews relevant psycholinguistic research into reading, particularly reading of Chinese text. This is followed by reviews of relevant research into translation as a reading activity, and a discussion of its role in language teaching and testing.

The core of the thesis is the discussion of the zero pronoun in Chinese, including discussion of anaphoric choice - the writer's decision on when to use zero in preference to an explicit anaphoric form - and of anaphoric resolution - how a reader decides what a zero pronoun refers to. Anaphoric resolution may be problematic for less experienced readers of Chinese owing to its lack of rich morphological inflection which, in other languages, provides the reader with information. Some of the key ideas on anaphoric choice and resolution are then applied to the analysis of the data in the parallel translations. It would appear that factors in Chinese texts which have an effect on comprehending zero pronouns are antecedent distance, topic persistence, abstraction, multiplicity of arguments and the meaning of the verb. Characteristics of the reader which may affect comprehension of the zero pronoun include personal schemata which may lead to elaborative inferences. On the basis of the data I suggest that mark schemes could be devised on a scalar system encompassing optimal solution, proximal solution and non-solution, which might help to solve the problem of variability in marking translation.

A by-product of the thesis, and an avenue for further research, is the apparent close relationship between idea units, clause length, punctuation breaks and antecedent distance in Chinese texts and saccade length and working memory capacity in the reader of Chinese.

## Table of Contents

### Chapter 0: Introduction

|  |   |
|--|---|
| 0.1 Aims of the thesis   | 1 |
| 0.1.1 Primary aim: investigation of the translation of the zero pronoun in Chinese into appropriate pronouns in English as evidence of reading processes | 1 |
| 0.1.2 Translation as reading and understanding   | 2 |
| 0.1.3 Use of parallel translations as data   | 2 |
| 0.1.4 Theoretical standpoints of the thesis  | 3 |
| 0.1.4.1 Interactive model of reading   | 3 |
| 0.1.4.2 Descriptive model of translation   | 3 |
| 0.1.4.3 Reading and translation of the zero pronoun  | 3 |
| 0.1.5 Secondary aim: evidence of reading comprehension in translation  | 4 |
| 0.2 The subjects in the study  | 5 |
| 0.2.1 Reading language   | 6 |
| 0.3 Examination conditions   | 6 |
| 0.4 Reservations about using translation as reading data   | 6 |
| 0.5 Theoretical basis for the study  | 7 |
| 0.6 The problem of orthography   | 8 |
| 0.6.1 Similarities and differences in Chinese reading and English reading  | 8 |
| 0.7 Structure of the thesis  | 9 |

## **Chapter 1: Psycholinguistic models of reading and their application to the reading of Chinese text.**

|  |    |
|--|----|
| 1.0 Introduction   | 13 |
| 1.1 Teaching English reading and teaching Chinese reading                                  | 13 |
| 1.2 The insights of Huey   | 14 |
| 1.3 Some models of reading -'bottom up,' ,top down,' and interactive                       | 17 |
| 1.3.1 The 'bottom up' model of reading   | 17 |
| 1.3.2 'Top down' models of reading   | 18 |
| 1.3.3 Interactive models of reading  | 19 |
| 1.4 Processing at various levels of reading in English                                     | 20 |
| 1.4.1 Phonological and visual access   | 21 |
| 1.4.2 Lexical access   | 22 |
| 1.4.3 The processing of function words in English  | 22 |
| 1.4.4 Processing of sentences  | 24 |
| 1.4.5 Constructing meaning   | 25 |
| 1.4.6 Propositional representation   | 25 |
| 1.4.6.1 Arguments and propositions within a framework of macrostructure and microstructure | 26 |
| 1.4.7 The role of schema in reading  | 27 |
| 1.4.7.1 Formal schemata  | 28 |
| 1.4.7.2 Content schemata   | 29 |
| 1.4.8 Inference  | 29 |
| 1.4.8.1 Memory and inference   | 30 |
| 1.5 Chafe's 'focuses of consciousness and spurts of language'                              | 32 |
| 1.6 Mental spaces  | 33 |
| 1.7 Summary and conclusion   | 37 |
| 1.7.1 A note on the usefulness of experimental studies of reading                          | 38 |

## Chapter 2: Chinese reading: traditional approaches, characteristic features of Chinese text and recent Chinese psycholinguistic studies of reading

|           |   |    |
|-----------|---|----|
| 2.0       | Aim and introduction  | 40 |
| 2.1       | About the Chinese language  | 41 |
| 2.1.1     | Chinese 'grammar'   | 41 |
| 2.1.2     | Chinese language and writing  | 43 |
| 2.2       | Chinese language and short term memory  | 44 |
| 2.3       | Chinese language and reading disability   | 46 |
| 2.4       | Levels or stages in interactive reading as applied to Chinese                                 | 49 |
| 2.4.1     | Some perceptual aspects of reading Chinese text - density, complexity, regularity and spacing | 49 |
| 2.4.2     | The psychological reality of the Chinese word - a word or a constituent?                      | 52 |
| 2.4.2.1   | <i>chengyu</i> as words   | 54 |
| 2.5       | Eye movement studies of Chinese reading   | 55 |
| 2.5.1.    | A note on scanning direction in Chinese reading   | 55 |
| 2.5.2     | Saccades and fixation and their relationship to reading comprehension                         | 56 |
| 2.5.3     | Moving window presentation  | 57 |
| 2.5.4     | Preferred landing position  | 57 |
| 2.5.5     | Progressive and regressive saccades and skipping in Chinese reading                           | 59 |
| 2.5.6     | Saccades, fixation and linguistic information   | 60 |
| 2.5.7     | Characteristics of Chinese text which may affect reading at syntactic level                   | 62 |
| 2.5.7.1   | Empty words and full words  | 63 |
| 2.5.7.2   | Pronouns and null arguments   | 66 |
| 2.5.7.3   | Lack of inflection in Chinese   | 66 |
| 2.5.7.4   | A note on punctuation   | 66 |
| 2.6       | Potential problems of comprehension of Chinese text   | 68 |
| 2.6.1     | Potential problems of low level decoding of Chinese text                                      | 68 |
| 2.6.2     | Grammatical/syntactic sensitivity in reading Chinese text                                     | 68 |
| 2.6.3     | Empirical studies of grammatical/syntactic sensitivity in reading Chinese text                | 69 |
| 2.6.3.1   | Timing and comprehension  | 69 |
| 2.6.3.1.1 | The role of the verb in comprehension   | 71 |
| 2.6.3.1.2 | The role of the noun and classifier in comprehension  | 72 |
| 2.6.3.2   | Cue validity or information value of linguistic forms in reading                              | 75 |
| 2.6.3.2.1 | Animacy and word order as cues in comprehension   | 75 |

|   |    |
|---|----|
| 2.6.3.2.2. Syntactic and semantic information as cues in understanding pronouns           | 78 |
| 2.6.3.3 Linguistic information affecting reading for comprehension and reading for recall | 82 |
| 2.6.3.3.1 Linguistic information affecting Chinese and English reading                    | 83 |
| 2.6.3.3.2 Thematic information in comprehension   | 85 |
| 2.6.3.4 The application of formal schema in reading Chinese                               | 87 |
| 2.7 Summary and conclusion  | 90 |

## Chapter Three: translation as reading

|   |     |
|---|-----|
| 3.0 Aim and introduction  | 94  |
| 3.1 Defining translation  | 94  |
| 3.2 Theories of translation   | 95  |
| 3.2.1. 'Bottom up' and 'top down' approaches to the theory<br>of translation; prescriptive and descriptive views              | 95  |
| 3.2.2 The sourcier (bottom up) approach to translation  | 96  |
| 3.2.2.1. The notion of equivalence in sourcier<br>approaches  | 98  |
| 3.2.3 The cibliste (top down) approach to translation   | 98  |
| 3.3 Descriptive views of translation  | 99  |
| 3.4 Recognising cognitive processes in translation -<br>la saisie du sens hors langue   | 100 |
| 3.4.1 Means of providing evidence of cognitive<br>processes in translation  | 101 |
| 3.4.1.1 Translation errors  | 101 |
| 3.4.1.2 Innate or natural translation   | 102 |
| 3.4.1.3 Investigation of children's natural<br>translation as a developmental phenomenon                                      | 102 |
| 3.4.1.4 Think-aloud protocols   | 103 |
| 3.4.1.5 Diachronic individual analysis  | 105 |
| 3.4.1.6 Translation corpora   | 105 |
| 3.4.1.7 Synchronic analysis   | 105 |
| 3.4.1.7.1 Translation norms: a descriptive<br>account of synchronic parallel translations                                     | 105 |
| 3.4.1.7.2 Hofstadter's parallel translations of<br>Clément Marot  | 107 |
| 3.4.2 A note on automatisisation or proceduralisation in<br>translation   | 108 |
| 3.4.3 A note on cultural and chronological factors in<br>translation  | 109 |
| 3.5 Translation into and from Chinese   | 110 |
| 3.5.1 Generalisations about the translatability of Chinese:<br>ellipsis, implicitness, inexpressibility and ideation          | 111 |
| 3.5.1.1 Extra-textual meaning in Chinese text: ellipsis<br>and implicitness   | 111 |
| 3.5.1.2 Extra-textual meaning in Chinese text:<br>inexpressibility  | 112 |
| 3.5.1.3 Extra-textual meaning in Chinese text:<br>direct ideation   | 114 |
| 3.5.2 The practical implications of ellipsis, implicitness,<br>inexpressibility and ideation for the translator<br>of Chinese | 116 |
| 3.5.2.1 Explication in translation  | 116 |
| 3.5.3 Chinese approaches to translation   | 118 |

|   |     |
|---|-----|
| 3.5.3.1 Fidelity, understanding and elegance: Yan Fu's translations of Western literature and science | 118 |
| 3.5.3.2 Lin Yutang: 'speaking for an old friend'  | 119 |
| 3.5.4 Some more recent defining approaches to Chinese translation                                     | 120 |
| 3.5.4.1 Huang Xuanfan: a <i>sourcier</i> to <i>cibliste</i> cline                                     | 120 |
| 3.5.4.2 Xu Suchong: expanding on <i>xin, da, ya</i>   | 121 |
| 3.5.4.3 Qian Gechuan: coherence and intelligibility   | 122 |
| 3.6 Attitudes to translation  | 123 |
| 3.6.1 Attitudes to translation in language teaching   | 123 |
| 3.6.2 Translation in language teaching and language testing   | 126 |
| 3.6.2.1 Arguments against the use of translation in language teaching and testing                     | 127 |
| 3.6.2.2 Arguments for the use of translation in language teaching and testing                         | 130 |
| 3.6.2.3 Problems with question types in reading comprehension exercises - can translation help?       | 132 |
| 3.6.3 Translation as a testing tool   | 134 |
| 3.6.3.1 The anomaly of measurement and authenticity   | 135 |
| 3.6.3.2 Testing communicative competence by means of translation                                      | 136 |
| 3.6.3.3 Testing strategic competence by means of translation  | 137 |
| 3.6.3.4 Marking translation and criteria for correctness  | 138 |
| 3.6.3.4.1 A further exploration of marking translation  | 139 |
| 3.6.3.4.2 Application of marking schemes for translation  | 141 |
| 3.7 Summary and conclusion  | 141 |

## Chapter Four: The role of the null argument or zero pronoun in Chinese discourse

|           |  |     |
|-----------|--|-----|
| 4.0       | Aim and introduction   | 143 |
| 4.1       | Null arguments, zero pronouns and anaphora across languages                                | 143 |
| 4.1.1     | A note on pro-drop languages   | 144 |
| 4.1.2     | Patterns of reference  | 145 |
| 4.2       | Characteristics of Chinese grammar which affect choice and comprehension of null arguments | 147 |
| 4.2.1     | Morphological structure of Chinese   | 147 |
| 4.2.1.1   | A note on redundancy   | 148 |
| 4.2.1.2   | A note on verblessness   | 148 |
| 4.2.1.3   | Deletion and transposition in Chinese  | 149 |
| 4.2.2     | The notion of topic in discourse in languages other than Chinese                           | 151 |
| 4.2.2.1   | Topic-comment structure in Chinese in both written and spoken texts                        | 154 |
| 4.2.2.2   | Topic-comment as syntax  | 156 |
| 4.2.2.3   | The problem of zero topic  | 158 |
| 4.3       | Factors affecting speakers' or writers' choice of full NP, pronoun or zero pronoun         | 159 |
| 4.3.1     | Anaphoric choice in languages other than Chinese   | 159 |
| 4.3.1.1   | Government and binding theory approaches to anaphora in languages other than Chinese       | 159 |
| 4.3.1.2   | Discourse based approaches to anaphora in languages other than Chinese                     | 160 |
| 4.3.1.2.1 | Rhetorical structure analysis as applied to anaphora                                       | 161 |
| 4.3.1.2.2 | Rhetorical structure analysis as applied to anaphora in English                            | 162 |
| 4.3.1.2.3 | Return pop in English  | 162 |
| 4.3.1.2.4 | Possible differences in anaphoric choice in conversational and written genres              | 163 |
| 4.3.1.2.5 | A discourse based view of anaphora in Japanese   | 163 |
| 4.3.1.2.6 | A compromise view of anaphora in English   | 165 |
| 4.3.2     | Anaphoric choice in Chinese  | 166 |
| 4.3.2.1   | Government and binding theory approaches to anaphora in Chinese                            | 166 |
| 4.3.2.2   | Discourse based approaches to anaphora in Chinese  | 167 |
| 4.3.2.2.1 | Relative frequency of nouns, pronouns and zero pronouns in Chinese                         | 169 |
| 4.3.2.2.2 | A centering based refutation of Li and Thompson's discourse analysis view                  | 170 |
| 4.3.2.2.3 | Topicality hierarchy in reference tracking   | 173 |



|   |     |
|---|-----|
| 4.3.2.2.4 Negligibility   | 176 |
| 4.3.2.2.5 Pronouns in a topicality hierarchy in<br>reference tracking   | 177 |
| 4.3.2.2.6 Noun anaphora in topicality hierarchy<br>in reference tracking  | 177 |
| 4.3.2.2.7 A Gricean account of zero and pronominal<br>anaphora in Chinese   | 178 |
| 4.3.2.2.8 A rhetorical structure analysis of<br>anaphora in Chinese   | 181 |
| 4.4 Anaphoric resolution  | 184 |
| 4.4.1 A study of language specific discourse strategies<br>in the processing of zero anaphors by Chinese readers<br>and English readers | 184 |
| 4.4.2 A study of anaphoric resolution in Chinese speaking<br>children and adults  | 187 |
| 4.5 Summary and conclusion  | 190 |

## Chapter 5: Analysis of the data: interpretations of zero anaphora in four short texts

|           |  |     |
|-----------|--|-----|
| 5.0       | Introduction   | 193 |
| 5.1       | The subjects   | 193 |
| 5.2       | The study  | 194 |
| 5.3       | Procedures   | 196 |
| 5.4       | Raters   | 198 |
| 5.5       | The texts  | 198 |
| 5.5.0     | Terms used in the analysis   | 200 |
| 5.5.1     | Text 1   | 203 |
| 5.5.1.1   | Reference chains and topic continuity in Text 1  | 204 |
| 5.5.1.2   | Subjects' anaphoric resolution of zero pronouns in Text 1  | 206 |
| 5.5.1.3   | Subjects' continuity of reference in Text 1  | 207 |
| 5.5.2     | Text 2   | 208 |
| 5.5.2.1   | Reference chains and topic continuity in Text 2  | 209 |
| 5.5.2.2   | Subjects' anaphoric resolution of zero pronouns in Text 2  | 211 |
| 5.5.2.3   | Subjects' continuity of reference in Text 2  | 213 |
| 5.5.3     | Text 3   | 214 |
| 5.5.3.1   | Reference chains and topic continuity in Text 3  | 215 |
| 5.5.3.2   | Subjects' anaphoric resolution of zero pronouns in Text 3  | 216 |
| 5.5.3.3   | Subjects' continuity of reference in Text 3  | 217 |
| 5.5.4     | Text 4   | 219 |
| 5.5.4.1   | Reference chains and topic continuity in Text 4  | 220 |
| 5.5.4.2   | Subjects' anaphoric resolution of zero anaphors in Text 4  | 221 |
| 5.5.4.3   | Subjects' continuity of reference in Text 4  | 222 |
| 5.6       | Summary of results   | 223 |
| 5.6.1     | The interdependence of syntactic and discoursal constraints on anaphoric choice in the texts used in the study | 223 |
| 5.6.2     | Subjects' anaphoric resolution shown in explicitation of zero pronouns   | 225 |
| 5.6.3     | Factors affecting reading comprehension  | 226 |
| 5.6.3.1   | Antecedent distance and proximity  | 227 |
| 5.6.3.2   | The positive effect of topic persistence   | 227 |
| 5.6.3.2.1 | Combined effects of topic persistence and antecedent distance  | 227 |
| 5.6.3.2.2 | The positive effect of persistence and continuity of topic   | 229 |

|   |     |
|---|-----|
| 5.6.3.2.3 The positive effect of context and/or<br>topic persistence  | 229 |
| 5.6.4 Theme and centrality  | 229 |
| 5.6.5 No apparent effects of animacy, gender, number,<br>or subject/object                                    | 230 |
| 5.6.6 Negative effect of abstraction and <i>chengyu</i>   | 230 |
| 5.7 Details of contributory effects not included in the tables  | 231 |
| 5.7.1 The negative effects of multiple arguments and topics not<br>distinguished by gender, animacy or number | 231 |
| 5.7.2 Negative and positive effects of verb meaning   | 232 |
| 5.7.3 The effect of schema  | 233 |
| 5.7.3.1 The importance of formal schema   | 233 |
| 5.7.3.2 The importance of content schema  | 233 |
| 5.7.4 The effect of elaborative inferences  | 234 |
| 5.7.5 The effect of return pop  | 235 |
| 5.7.6 Noteworthiness and punctuation  | 235 |
| 5.8 The problem of omissions  | 236 |
| 5.9 Conclusion  | 236 |
| 5.9.1 The combined effects of antecedent distance<br>and topic persistence on short term memory and cognition | 237 |
| 5.9.1.1 Reading and memory  | 238 |
| 5.9.1.2 Reading and saccades  | 239 |
| 5.9.1.3 Reading and idea units  | 239 |
| 5.9.1.4 Reading and punctuation breaks  | 240 |
| 5.9.1.5 Reading and mental spaces   | 241 |
| 5.10 Zero pronouns, memory, saccades, punctuation, idea units,<br>mental spaces and the reader                | 242 |

|  |         |
|--|---------|
| <b>Chapter 6: implications for further research</b>                            |         |
| 6.1 Revisiting the aims of the thesis  | 243     |
| 6.1.1 Primary aim  | 243     |
| 6.1.2 Secondary aim  | 243     |
| 6.2 The extent of the role of translation in analysing reading processes       | 244     |
| 6.3 What translation reveals about comprehending reference patterns in a text  | 245     |
| 6.3.1 Using translation as a reading comprehension exercise                    | 247     |
| 6.3.2 Parallel translations  | 247     |
| 6.3.3 Translation in language testing  | 248     |
| 6.4 The need for joined-up research into features of text and reader behaviour | 249     |
| 6.5 A summary of possible ways forward   | 250     |
| <br><b>References</b>  | <br>252 |
| <br><b>Appendices</b>  | <br>269 |
| Appendix 1: tables of subjects' actual anaphoric resolution                    | 270     |
| Appendix 2: raters' questionnaires   | 301     |
| Appendix 3: examination papers   | 314     |

## Acknowledgements

I would like to thank Edexcel Foundation for allowing me to use examination materials for my research, and the examination candidates for providing a rich seam of material for me to mine. I am indebted to Tony Howatt, my main supervisor, for guidance, reminders about coherence and relevance, and above all for his wisdom and perspicacity. I am also grateful to John Joseph for his unique assistance in the area of translation studies. My family, Mike, Anne and Andrew deserve praise for their tolerance of the twin tides of paper and jargon and their understanding of my physical absences and mental pre-occupations.

I would like to thank Kao Chung-yao and Stewart Paton for their sensitive and constructive comments on my manuscript at various stages in its production, and Linda Anderson for patient and sympathetic IT support. Patricia, Karin, Joseph, Chung-yao and Richard, all in TAAL, have offered moral support as and when needed, and David Brown, Jessica Marquis and Chau Manling, of Edexcel, have given me professional encouragement.

This thesis is dedicated to my parents, Herbert and Ivy Pellatt.

## Chapter 0: Introduction

### 0.1 Aims of the thesis

#### 0.1.1 Primary aim: investigation of translation of the zero pronoun in Chinese into appropriate pronouns in English as evidence of reading processes

The aim of this thesis is to investigate some of the processes of reading Chinese text. The investigation was carried out by comparing and analysing parallel translations of Chinese source texts into English target texts. The focus is on the interpretation, in these translations, of the null argument, or zero pronoun, in the Chinese language; Chinese is recognised as a 'pro-drop' language, and many researchers regard the zero pronoun as the 'normal mode of anaphora' in Chinese (Li and Thompson 1979). Whilst the data analysis focuses on a single aspect of Chinese - the zero pronoun and the ways in which it is comprehended - the review of literature undertaken is broad, necessarily encompassing reading, translation, and a review of characteristics of the Chinese language including a detailed account of the zero pronoun itself. In reviewing these distinct, but related areas, I hope to make connections which, until recently, have not usually been made, particularly in the relationship between reading, translation and language testing. The structure of the thesis encompasses a review of studies of reading as an interactive process in Chapter One and the particular case of reading Chinese in Chapter Two. In Chapter Three theories of translation, including Chinese theories of translation, are reviewed and translation as an outcome and as evidence of reading is discussed. Chapter Four focuses on the zero pronoun as a feature of Chinese which may be used to exemplify reading and translation processes. Finally, Chapter Five contains an analysis of translation data; these data illustrate reading comprehension by means of English interpretations of Chinese zero pronouns. Chapter Six comprises a discussion

of implications which may be drawn about the use of translation as a tool for testing reading comprehension.

### **0.1.2 Translation as reading and understanding**

My notion of translation is that it is not simply a process of matching 'equivalent' expressions in two languages; it is an activity to which the processes of reading and understanding are prior and indispensable. In the case of a very experienced translator who has automatised both languages, it may appear to be a simple matching operation, but in fact, the matching in the target language can only be done on the basis of a thorough reading of the source text, and a thorough comprehension of that reading. As I will explain further in Chapter 3, my view is that reading comprehension constitutes the major effort of the translation process as a whole. Even if more *time* is spent on the production end of the process, the comprehension end of the process governs the outcome.

### **0.1.3 Use of parallel translations as data**

I am not concerned with translation in polished, literary form, but as the expression of the source text's meaning in the reader's own language - language of habitual use or 'first' language (L1). The target texts used as data are spontaneous renditions made without the use of dictionaries or any other kind of consultation, and within a time limit.

As will be explained in the analysis of the data, I am not concerned with 'right' or 'wrong' answers, but with a range of possible translations and the reasons why those particular choices may have been made by the subjects. In the case of translating Chinese texts into English, inevitably, zero pronouns will be encountered, which in most cases, though not all, will need explicitation in English - that is, the addition of an overt anaphor, in the form of a NP or pronoun. This type of explicitation shows how the

reader/translator has understood the zero pronoun. The range of answers is taken as a pattern of responses across the group of subjects as a whole which may provide insights into the process of translation and hence the process of reading. This range of answers includes the most popular answers, which in most cases coincide with those regarded by adult raters as 'most accurate' and termed optimal solutions, and also includes less popular answers which may or may not be regarded as inaccurate, termed proximal solutions, and non-solutions.

#### **0.1.4 Theoretical standpoints of the thesis**

Two theoretical standpoints underpin the thesis. Firstly, that of an interactive model of reading, in which the reader resolves zero anaphora by means of both syntactic and pragmatic criteria, and secondly, a descriptive model of translation, to which this model of interactive reading is central.

##### **0.1.4.1 Interactive model of reading**

An interactive model of reading, as proposed by, for example Rumelhart (1981) and Grabe (1988), assumes parallel and simultaneous processing of a text at all levels: lexical, grammatical (morphological), syntactic and schematic.

##### **0.1.4.2 Descriptive model of translation**

A descriptive model of translation investigates how translation is done rather than how it should be done, focusing on parallel translations of the same source text by a number of translators, as exemplified by the work of Toury (1995).

##### **0.1.4.3 Reading and translation of the zero pronoun**



Within the framework of the interactive model of reading I aim to focus on comprehension and translation of the null argument, or zero pronoun, for the following reasons:

1. The null argument, or zero pronoun, in Chinese has been well researched, both as argument and as anaphor<sup>1</sup>, but mainly using material which is predictably ambiguous. However, in my data I have found instances where readers' interpretations differ, but where ambiguity would not normally be expected.
2. Various arguments have been put forward for anaphoric choice (why and how a speaker or writer selects a zero pronoun) and anaphoric resolution (how a listener or reader decides what that zero pronoun refers to). These may include criteria of proximity, topic continuity, animacy, meaning of the verb and centrality.

A group of reader/translators' varied interpretations are analysed in order to discover what cues they may have used to determine anaphoric resolution. The key question is: how have the subjects in the study used other features of the text to decide who does what to whom when there is no overt pronoun to provide this information?

### **0.1.5 Secondary aim - evidence of reading comprehension in translation**

A secondary aim is to show how translation gives evidence of comprehension and may therefore be a useful teaching and assessment tool. Where readers are required to translate zero pronouns in a Chinese text into non-zero pronouns, or noun phrases, in English, the resulting target text may serve to illustrate interaction between levels in an interactive model of reading. In this

---

<sup>1</sup> The term 'null argument or zero pronoun' is used throughout Chapters 0, 1, 2 and 3 of this study so as to include non-anaphoric instances of zero, but from Chapter 4 on, only the term 'zero pronoun' is used.

way, contrary to current received wisdom in language teaching circles, occasional translation of carefully selected texts could contribute to the teaching and testing of reading comprehension.

## **0.2 The subjects in the study**

The subjects in the study are groups of bilingual Chinese/English readers of varying abilities who translated Chinese texts into English for the purposes of an examination. The study focuses on three texts from the 1998 Edexcel London Examinations 'A' Level paper and one text from the 2000 examination. On each occasion approximately 100 subjects' translations were analysed. The subjects who took the 2000 examination were different from those who took the 1998 examination, but were of similar age and from the same range of educational centres.

Given the source of the data, it was impossible to control for language ability, age, or gender. However, the fact that the rubrics of the examination are in English indicates that a certain standard of English literacy is required. For most of the subjects English is the language of habitual use. Though some subjects showed evidence of weakness in one or other language, and some (probably from Malaysian or Vietnamese backgrounds) in both, they were all bilingual to a greater or lesser extent. Those translations in which the English or the Chinese was so weak as to be illegible, unmarkable or incomprehensible were omitted from the study. For the purposes of this study, English is regarded as L1, Chinese as L2.

Subjects originate from various Chinese speaking areas - Mandarin speaking areas of the People's Republic of China; the Cantonese speaking areas of Canton and Hong Kong; Macao, Taiwan and the diaspora in South Africa and the UK and other parts of Europe. The translations used in the study were all from the 'home' (ie British) centres, thus all the subjects in the study

were attending British schools or colleges at the time of the examination. For most of these young people, English is the language of habitual use, and I am therefore regarding Chinese as L2 and English as L1 for the purposes of the study. The subjects were all likely to be aged between 16 and 19, though a tiny minority of the subjects were mature students who took the examination at colleges of further education.

### **0.2.1 Reading language**

The subjects are required to read Modern Standard Chinese, which is most closely associated with putonghua, the standard Chinese used in education and media in mainland China. Speakers from other dialect areas within the PRC, Cantonese speakers from Hong Kong, and 'Guoyu' speakers from Taiwan are all taught to read this type of written Chinese even if they do not speak putonghua. The examination offers candidates the choice between full characters, as used in Taiwan and Hong Kong, and simplified characters, as used in the mainland PRC and Singapore. The texts used in the study are presented in simplified characters in this thesis.

### **0.3 Examination conditions**

While these translations were not made in experimental conditions, there is a degree of uniformity imposed by the examination situation. The time spent on the texts is limited, subjects are not permitted to consult a dictionary and the tasks are carried out in conditions of silence.

### **0.4 Reservations about using translation as reading data**

It could be argued that translation into L1 is evidence of production of L1 rather than evidence of understanding of L2. At sophisticated levels of translation this may be true, but at the level at which I am investigating, it is likely that only a very few subjects are in a position to manipulate their interpretation of the text beyond certain limits. In addition, critics may argue

that the range of ability covered in this study is too wide. Unfortunately the nature of the data prevents controlling for ability, age or gender. However, the subjects are self-selecting to a certain extent, as the examination can only be taken on the basis of a certain level of literacy in both English and Chinese. Target texts which were not completed, or which were incomprehensible owing to lack of literacy skills in English, thus providing insufficient information about the candidate's understanding, were omitted from the study. Some teachers and testers would be opposed to bilingual testing or teaching on ideological grounds. I myself subscribe to the ideal of target language teaching, but I feel that there is value in bilingual approaches in testing and monitoring, a view also held by, among others, Källkvist (1998), Lu (2000), Malmkjaer (1998), Stibbard (1998), Yan (2000), and Yu (2000). This view is more fully defended in Section 3.6.

### **0.5 Theoretical basis for the study**

The thesis assumes reading comprehension to be a central part of translation and thus views translation as a product which shows overtly what and how the reader/translator has comprehended. The study also assumes an interactive model of reading, that is, a process in which items at various levels of linguistic and non-linguistic information are processed simultaneously, each triggering or triggered by and assisting or assisted by other items.

Various approaches to reading, broadly including bottom-up models, top-down models and interactive models are discussed in Chapter 1 with reference to English. The reason for this is two-fold: there has been much research into the reading of English and this research provides a psycholinguistic framework against which to assess the body of research into the reading of Chinese; furthermore, much research into the reading of Chinese, particularly experimental work, has been based on or inspired by the

research into English, or has directly replicated experiments and studies on English.

## **0.6 The problem of orthography**

There has been much misunderstanding of the way in which Chinese reading proceeds. This is largely due to the misconception of the nature of Chinese orthography. A great deal of research into Chinese reading has focused on the physical characteristics of the orthography, particularly its apparent predominantly 'visual' nature. Great emphasis has been placed on the perception of print, ignoring the characteristics of Chinese text and the importance of language in reading.

### **0.6.1 Similarities and differences in Chinese reading and English reading**

The processing of print seems not to differ greatly as between English and Chinese, both languages having a great deal of orthographic irregularity, but both having some systematic rules of 'spelling,' and it is likely that readers of both access words or characters both visually and phonologically. Any differences in the reading process are likely to arise from the grammatical and discoursal structure of the language rather than the characteristics of the orthography. One of the chief differences is the high incidence of pronouns used in English which make anaphoric reference relatively explicit and the contrasting tendency towards zero pronouns in Chinese, which creates a rather more implicit pattern of reference.

A description of the Chinese language is crucial to an understanding of translations into English. One of the main features of succinctness in Chinese is the null argument or zero pronoun. Thus, there are situations in which a linguistic item which in English would be regarded as a proposition, in Chinese consists of predicate only, a factor which may affect reading

comprehension. In these cases translation into English from Chinese may necessitate explicitation, that is, the addition of items not apparent in the source text, which make clear the reader/translator's comprehension.

The wide range of anaphoric resolutions made by bilingual Chinese/English reader/translators shows, firstly, what options are available for anaphoric resolution when a reader is confronted by a zero pronoun, and secondly, how the decision is made. Readers may decide who does what to whom by using syntactic cues such as proximity, discursual cues such as topic continuity, pragmatic cues such as animacy, semantic cues such as the meaning of the verb or by using their own background knowledge.

## **0.7 Structure of the thesis**

The scope of the thesis is necessarily broad, including reading theory, translation theory and grammar and discourse. I hope to make overt connections which are usually ignored - mainly, the crucial, close relationship between reading and translation and secondly, the usefulness of translation in assessing reading comprehension.

**Chapter One** As so much Chinese reading research has been inspired by, or has directly replicated, studies of English reading, Chapter One summarises approaches to the reading of English, including theoretical and descriptive models and prescriptive or pedagogical principles. This includes a brief description of bottom-up and top-down models and the interactive model which brings together strands of both. Bottom-up models emphasise the importance of decoding at orthographic level and word recognition. In the classroom this would translate into an emphasis on phonics. Top-down models emphasise the importance of readers' general knowledge and hypotheses they make and test as they read. Interactive models take the perhaps common sense view that reading consists of far more than sounding out words on the one hand or guessing the plot on the other. Chapter One

should serve as a background against which to assess the studies of Chinese reading reviewed in Chapter Two.

**Chapter Two** Chapter Two begins with a very brief description of some characteristic properties of Chinese language and writing which may affect reading, particularly its tendency to succinctness. This is followed by discussion of processing of Chinese at grammatical and discoursal levels. A number of experimental studies are reviewed which replicate studies of English reading and which tackle the comprehension of Chinese text at various levels, from character level through sentence level to paragraph level.

**Chapter Three** This chapter reviews Western and Chinese theories of translation and justifies the use of translation data as evidence of the reading process. Reading and understanding are major parts of the translation process and are evident in the product, the target text. Approaches to translation are diverse and, as in reading, there are differences between top-down and bottom-up, descriptive and prescriptive approaches.

For the purpose of this thesis I have adopted a descriptive approach inspired by that of Toury (1995), who has investigated the use of and violation of norms in parallel translations by student translators.

While the first part of the chapter deals with translation in general, highlighting the role of comprehension in the translation process, the second part of the chapter is devoted to translation to and from Chinese. The characteristics of Chinese text which might create difficulties in reading, particularly its tendency towards ellipsis, are explored, through the theoretical or ideological stances of some representative translators of Chinese.



My stance on translation as evidence of reading comprehension is likely to be subject to criticism, as translation has so long been in disfavour in language teaching classrooms. As a newly qualified teacher in the 1970s, and throughout most of my teaching career, I myself espoused the anti-translation lobby. My recent experience as a tester has caused me to think otherwise. This thesis does not provide a platform for a blow by blow defence of translation in the classroom, but I wish to show that it may have some advantages in the area of teaching and testing comprehension of the target language. My emphasis, therefore, is on translation from L2 to L1.

**Chapter Four** This chapter takes up the notion of the ‘silences’ - the elliptical nature of Chinese - and in particular the null argument or zero pronoun. The topic-comment structure in Chinese, which is closely related to the zero pronoun, is discussed. Anaphoric choice, that is, the writer’s decision as to which type of referent to use, and anaphoric resolution, that is, the reader’s decision as to what the anaphor refers to, are discussed. Anaphoric resolution is central to this thesis, and a number of studies are reviewed which reveal criteria or cues used in readers’ judgements. These include both discourse analysis approaches and syntactic approaches, encompassing criteria of proximity, topic continuity, animacy, word order and the meaning of the verb.

**Chapter Five** Presentation of data. Diagrams of the four Chinese source texts are presented showing syntactic and discoursal patterning of the zero pronoun. These are followed by tables showing percentages of subjects opting for optimal, proximal and non-solutions in relation to certain features of the texts - antecedent distance, topic persistence, abstraction, animacy, gender, number and subject/object roles. Where numbers of optimal solutions are high, it may be that certain features of the text help the reader, and where optimal solutions are low and proximal and non-solutions are



high, certain features of the text may disrupt the reader. There are also tables showing subjects' continuity of reference. The results are discussed, first in relation to the broad outcome, which focuses on the importance of topic persistence, and then in relation to details of factors underlying non-solution.

**Chapter Six: Conclusions and implications** The findings of the study are discussed in the way in which they relate to other reading research on short term memory capacity, eye movements, idea units and mental spaces. Some suggestions are made for further research.

## **Chapter 1: Psycholinguistic models of reading and their application to the reading of English and Chinese text**

### **1.0 Introduction**

This chapter reviews some Western approaches to reading, which may be useful as background to recent studies of Chinese reading. Section 1.2 is a review of the classic studies of Huey (1908) which are still of relevance today. This is followed by a review of studies which have highlighted eye movement, for example those by Rayner and Pollatsek (1989). Research on phenomena such as fixation and saccades, though far from giving a complete picture of reading processes, sheds light on syntactic, semantic and discursal processing. Eye movement studies of Chinese reading have replicated studies of English reading at various levels beyond automatised processing of print. In addition to eye movement studies, research using recall and comprehension questions is also reviewed. The focus of the reviews is on studies of syntactic processing and meaning construction.

### **1.1 Teaching English reading and teaching Chinese reading**

The teaching of native speaker reading, in the UK, has drawn on the use of both phonics and whole word teaching, owing to the irregularity of English spelling. In the teaching of Chinese reading, the emphasis has always been on the recognition of the single character, its pronunciation and meaning, taught and learned holistically. This method was (and still is) perceived by many as visual rather than phonological. Radicals, or meaning components of Chinese characters, (see Section 2.1.2) are taught, but it seems that the approach is not always systematic.<sup>2</sup> Children would normally learn individual characters and chunks of text by rote memorisation. Both Chinese and English are very

---

<sup>2</sup> Kang Baowen complains that in Hong Kong primary schools characters are not taught in order of difficulty, or by radical, or by phonetic family. He suggests teaching in phonetic families, in ascending order of difficulty and using the radicals to explain semantic structure. (Kang 1999) This is a very rational approach, but takes into account neither character frequency nor subject matter. The target of

irregular in their orthographic phonological correspondence. However, for the purposes of this thesis it is assumed that reading processes at the level of perception of print are automatised and will not be covered in any detail, except where miscues at this level affect processing at other levels. The focus of the thesis is on processes at syntactic and discoursal level, particularly interpretation of the zero pronoun.

## **1.2 The insights of Huey**

In any review of modern approaches to reading, it is worth looking back at the work of Huey (1908), for, although writing at the beginning of the twentieth century, he addresses the issues which are still hot topics of debate at the beginning of the twenty first century.

Huey reviews a number of contemporary theories and methods of teaching reading (mainly English, but also some European languages) and describes in some detail experimental research on reading of the 1890s and early 1900s. Huey's motivation is that 'we need to know just what a child does when he reads' in order to devise good teaching methods (Huey 1908: 9). Later generations have re-visited the material reviewed by Huey, and most of it now seems eminently sensible, or even taken for granted, in particular Huey's opposition to the teaching of letter names, and the principle that 'reading may and must go on by other means than the recognition of letters' (Huey 1908: 68). For example he reviews an early example of research on word superiority effect, a phenomenon covered in more detail in Section 2.4.2 on the psychological reality of the Chinese word. Experiments such as this, and eye movement studies, are still being replicated and developed, particularly in association with Chinese reading (see Section 2.5). More recent research

---

his criticism, the character list currently in use in schools in Hong Kong, is based on frequency and applied in a situational approach. Chen (1997) found that children learning characters by means of CAL learned more effectively when the pronunciation of the character was provided simultaneously with the visual presentation.

has found eye movement studies, and their off-shoot 'moving window' studies, to show a great deal about parsing, prediction and use of context in decoding, for example, in work by Rayner and Pollatsek (1989), Chen (1999) and Just and Carpenter (1976) and about focusing of consciousness (Chafe 1980).

Among the conclusions reached by Huey is the importance of 'reading for thought' and what emerges from his book is that for children meaningfulness is the key to understanding (Huey 1908: 125). He attaches some importance to apperception - 'some material is apperceived and the rest is associatively supplied' (Huey 1908: 84). He categorises readers as two different types, the subjective and the objective. The subjective type has 'wandering attention,' 'slight fidelity,' and a 'large associative contribution,' apperceiving words from the total character of the word form, rather than the dominating parts, for example, recognising letters which protrude above and below the line, but misreading small letters and vowels (Huey 1908: 92). In this 'subjective reader' perhaps there are signs of what would now be thought of as dyslexia. Huey's objective reader, on the other hand, recognises the dominant parts of the word, reads a smaller amount at a glance, but is less liable to error. So Huey is saying that careful readers are better readers, perhaps thinking in line with the author rather than thinking along their own lines, but he also insists that 'the manner of perceiving words must depend, for the child, very largely on how he is taught to perceive them in learning to read' (Huey 1908: 103).

There are points on which modern teachers would perhaps disagree with Huey, for example he observes hindrance in spacing between words rather than help, suggesting that printing has brought about the division on paper of what belongs together in speech (Huey 1908: 115). This is debatable: spaced writing has developed from non-spaced writing, presumably for common sense reasons. Certainly, in the case of Chinese writing, where there

is equal (tantamount to no) spacing between characters, but very noticeable spacing where punctuation occurs, learner readers and children may have problems perceiving boundaries between words, but perhaps not between clauses (see Sections 2.4.1 and 2.5.7.4).

Huey has some interesting things to say about reading, meaning and translation:

*'that the general meaning dawns upon the reader precedent to the full sentence-utterance is evidenced by many cases in which variant words of equivalent meaning are read, and also by the comparative ease with which a reader may paraphrase the thought of what he reads. This is especially noticeable in the case of a person reading a foreign language which he does not pronounce easily, but which he comprehends rather rapidly. Here the visual word and phrase percepts touch off total meanings which clothe themselves, as the meanings become articulate, in English sentences, and we have as a result the mongrel reading which passes for French or German in so many modern language classes' (Huey 1908: 147).*

This observation firstly relates to the way in which a reader brings his or her own background knowledge and schemata to bear on a text which he may not have perfectly understood. It also shows how paraphrase in the L1 can illustrate the processes of understanding and Huey's remark about 'mongrel reading' is somewhat reminiscent of Nabokov's (1955) differing view that a 'schoolboy's boner is less of a mockery ... than commercial interpretation' (Nabokov 1955). The relationship of translation to reading is more fully discussed in Chapter Three.

Huey shows how context affects reading - the more information is supplied in the context, the more specific and close the interpretation is (Huey 1908: 156), a notion which is covered throughout Sections 1.4, 1.5 and 2.6.

Huey addresses the issue of function words. He reviews, but does not comment on, Dearborn's (no date) suggestion that 'connective and non-substantive words make the greatest demands on perception and thus require

most fixations' (Huey 1908: 47). Dearborn, like Yang and McConkie (1999) investigating Chinese readers nearly 100 years later, found that English readers fixated anywhere, including the gaps between the words. Perhaps Dearborn's message was that readers *ought* to take note of grammatical function words, but in fact do not. More recent research has found that readers of both Chinese and English are more likely to fixate content words (see Section 2.5.6). In his own research Huey found that his subjects reported that while relational or function words suggested no imagery, they sometimes caused expectancy. Among these subjects, content words suggested a greater amount of imagery to some subjects than to others, implying variation in application of schema among individuals. The same subjects also remarked on the predictive quality of context - experiencing fulfilment or disappointment in expectation (Huey 1908: 157). The problem of function words in Chinese is addressed in Section 2.5.7.1. Huey's work is relevant here in that it adumbrates the interactive reading models of the late twentieth century, reviewed in Section 1.3.3.

### **1.3 Some models of reading - 'bottom up,' 'top down,' and interactive**

#### **1.3.1 The 'bottom up' model of reading**

When taught according to what are now known as 'bottom up' models, the reader was assumed to start with the very smallest unit of text - phonemes or letters in the case of European languages - building up into progressively larger units of text such as words, phrases, clauses, sentences and groups of sentences. Mastering sound-symbol relationships was, and still is, considered to be of primary importance in developing reading proficiency. It is of relevance in teaching beginner reading of orthographies which have a high orthographic-phonemic correspondence.

### 1.3.2 'Top down' models of reading

In the West some researchers long promoted the idea that reading was also a matter of constructing the author's meaning through general and cultural knowledge but this view was not widely acknowledged until the 1970s. The fashions of the 1960s and 70s were for pupil centred, concept driven learning and during this time a top down approach, reflecting Huey's idea of reading for thought, emerged to challenge decades of what had appeared to be a narrow concentration on correct spelling and rote learning.

Goodman (1967), calling reading a 'psycholinguistic guessing game,' pioneered a reader centred view. His theory was that 'efficient readers minimise dependence on visual detail' (Goodman 1975). According to his model a competent reader neither needs to use, nor does use, all the information available in the text. Goodman's model inspired or provoked later theoretical models of reading which recognised that the reader needed to be actively involved in processes at multiple levels; necessarily processing print on the page, comprehending vocabulary and grammatical structures, retrieving information in the form of arguments and propositions, making inferences on the basis of the information provided and making predictions about the forthcoming text.

Goodman was concerned that 'language should be studied in process. Like a living organism it loses its essence if it is frozen or fragmented' (Goodman 1975). He felt that the emphasis on recognising words would fragment language and hamper good reading. While he admitted that reading must use visual input, he seemed to be hostile to any kind of spelling or phonological processing. He felt that the holding of an image in short term memory was necessary, but that the image should be of visual features of words rather than sound symbol correspondence. His theory of miscue analysis was based on performance errors of oral reading and the 'short circuits' (Goodman 1975)



that resulted from the miscues. According to Goodman miscues arise from letter naming (phonics), recoding (generalisation - eg 'bean' must rhyme with 'mean'), syntactic nonsense (the reader generates 'an approximate deep structure without going beyond to meaning') and partial structures (the reader may sometimes read successfully, sometimes not). All of these he attributes to instruction (Goodman 1975), an idea which reflects Huey's (see Section 1.2). The theory had to apply to all languages, except in specific characteristics of syntax and orthography. He was probably correct in his view that instruction accounts for errors, as well as accuracy. Goodman's theory is in stark contrast to that of Adams' for example, whose pedagogic approach insists that context is irrelevant, and good reading is a strictly bottom-up phonic/spelling process (Adams, 1990).

Teaching methods are culturally based, and may change with time, as those in the English speaking world have done. Methods of teaching reading may be constrained on the one hand by the nature of the script and on the other hand by ideological, economic and social systems.

### **1.3.3 Interactive models of reading**

Goodman's model paved the way for models which allowed for both careful decoding at orthographic level and for processes which infer the author's intended meaning and take account of the reader's general knowledge.

The models of the 1970s, including that of Goodman, tended to be linear. Linear models are deficient in that they do not take account of some processes which are known to take place in reading, such as the effect of syntax on word perception (Samuels and Kamil 1988: 27,28). Early interactive models such as that of Rumelhart (1977), in which information from syntactic, semantic, lexical and orthographic sources converge (Rumelhart 1977: 589-590) took account of possible simultaneous processing of automatised decoding at



letter and word level (or in the case of Chinese and Japanese, syllable level), lexical access, parsing skills, knowledge of text structure (formal schemata) and background knowledge (content schemata). Researchers like Stanovich (1980) reacted to both linear bottom up models which allowed for no feedback or monitoring stages, and top down models in which reading is conceptually driven. Stanovich's main criticism of the top down models was that only weak readers would use such a model (a view which reflects Huey's description of subjective readers, see Section 1.2), as skilled readers would recognise words without difficulty and would not need to make predictions. Stanovich incorporated aspects of skilled and unskilled reading into his model so that 'a process at any level can compensate for deficiencies at any other level' (Stanovich 1980: 36) Grabe (1988) points out that an interactive model includes any which assumes parallel processes and accounts for more than simply linear processing (Grabe 1988: 60). Many of these models are based on empirical studies and may be regarded as descriptive - they describe how reading *is* done, rather than how reading *should be* done.

The different approaches cannot be strictly divided, and are not mutually exclusive. There are strong and weak forms of each, and there are elements common to each. A reliable model of reading is likely to encompass views and experimentation covering all possibilities. What is important is that models are built on evidence, and not vice versa. This thesis is based on the notion that reading involves interactive processing, evidence of which is found in parallel translations.

#### **1.4 Processing at various levels of reading in English**

The process of reading in English is not central to this thesis but needs to be discussed for two reasons: firstly, the bulk of reading research has been carried out on English reading and research on Chinese reading owes much to this body of work. Chinese researchers have used studies of English

reading as starting points, replicating and developing experimental and analytic methods and thereby discovering both similarities and differences. Secondly, it may be argued that English, compared with other European languages, bears a relatively greater resemblance to Chinese, in its linguistic structure, which includes much less morphological inflection, and particularly in its rather opaque orthography.<sup>3</sup> These characteristics of the languages will be explored more fully in later sections and chapters. The account of reading processes of English which follows is necessarily brief and focuses on particular aspects of the reading process which are relevant to this thesis.

The focal point of this thesis is the processing of the null argument in reading, and strictly speaking, 'lower' level decoding, such as word and character recognition and lexical access are not central to the argument. However, it is necessary to touch on these points, as differences at low levels of decoding may affect higher level processes. Secondly, assuming reading to be an interactive process, lower level decoding will necessarily affect higher level decoding.

#### **1.4.1 Phonological and visual access**

There has been a long, intense and heated debate in the English reading world about the degree of phonological versus lexical access in reading individual words. Needless to say, there is support for and evidence of both, perhaps largely owing to the high degree of irregularity in English spelling. Rayner and Pollatsek (1989), whilst fiercely partisan in many aspects of their research, admit that while individual orthographies may steer readers to either the phonological route or the visual route, their own evidence suggests that 'both

---

<sup>3</sup> Inflection in English is minimal (no case endings except for pronouns, no gender or number agreements in verbs or adjectives), there is a penchant for ellipsis and there is a looseness of reference similar to that found in Chinese. English spelling is as irregular as Chinese 'spelling.'

routes are used in all languages.’ Not only that, but they also suggest that reading is a ‘culture free cognitive activity’ at the automatised level of perception of print (Rayner and Pollatsek 1989: 99). Orthographic decoding is not the central concern of this thesis and an automatised processing of print is assumed prior to processing at syntactic level and above.

#### **1.4.2 Lexical access**

There are obvious differences between processing of Chinese, in which the basic unit of processing in reading is the character (further discussed in Section 2.1.2) and English and other alphabetic systems in which the word is the basic unit of processing in reading, that is to say, the orthographic system provides spaces between these units. The main bone of contention in researching lexical access is the degree of contextual effect. Rayner and Pollatsek, for example, claim that ‘words are processed in much the same way in isolation as in text. While context somewhat affects the speed of processing words, its effects are surprisingly small’ (Rayner and Pollatsek 1989: 62). This is clearly a ‘bottom up’ view, and supporters of ‘top down’ and interactive theories would disagree. My own parallel translation data suggest that context is very powerful, providing not only semantic and syntactic clues to a correct reading, but sometimes also sending the reader on a wild goose chase.

#### **1.4.3 The processing of function words in English**

As is discussed in Section 2.5.7.1 function words play a crucial, though perhaps anomalous role in reading Chinese text. Does English reading research make any major contribution to our understanding of the processing of function words? Investigations of brain damaged patients have shown cases where, while content words are still produced and comprehended well, the processing of function words is seriously impaired, for example in the work of Coltheart, Patterson and Marshall (1980). Rayner and Pollatsek note, in relation to this, that readers do not extract information from the page

during saccades, but only during fixations (Rayner and Pollatsek 1989:113). Fixations are likely to be on content words rather than function words. This is where there could be differences in the reading of Chinese and English, and eye movement studies on Chinese reading are discussed in Section 2.5. An over simplified, but perhaps important observation to make at this point is that in English text content words are visually long and function words are visually short, thus ensuring that content words are highlighted and fixated: words which are short, frequent or predictable are most likely to be skipped, though it appears that the parafoveal preview provides predictive information. In Chinese text content words are visually short (though dense: see Section 2.4.1), but function words are even shorter, and often omitted, thus also ensuring highlighting of content words. Rayner and Pollatsek, on the basis of eye movement studies, go so far as to suggest that 'if a word is skipped it either must have been identified before it was skipped, or the reader simply made a guess as to what the word was without having seen it' (Rayner and Pollatsek 1989: 139). Noting that word skipping is a ubiquitous part of the eye movement record, they suggest that the parafoveal view, fairly far to the right of fixation, can provide the meaning of a word. It is fairly certain that a parafoveal view can facilitate later identification of a word or character.

The phenomenon of skipping of function words in Chinese is discussed in Section 2.5.7.1 and the perhaps related ubiquitous null argument or zero pronoun in Chinese text is discussed in Section 2.5.7.2 and throughout Chapter Four. My view is that the relationship of function and content words in Chinese, and the relationship of both to cognitive style is crucial to understanding the Chinese reading process.

#### 1.4.4 Processing of sentences

It should be noted here that the use of the sentence as a processing unit is a convenience. In English, capital letters and full stops neatly divide these units which provide handy material for experimenting and analysing. For reasons explained in 2.5.7.4, the sentence is not such a handy unit in the case of other languages, Chinese in particular.

Once words have been identified, how are they integrated into the text? Does the reader process meaning and analyse syntax independently, or simultaneously? According to Rayner and Pollatsek, syntactic parsing goes on more or less independently of processing of meaning. In cases of syntactic ambiguity, which are legion in initial parsing, as only parts of the sentence are available, it is likely that only one syntactic structure is computed by the reader, and only if it is incorrect is a new structure applied. Apparently, plausible meaning only has an effect on parsing when the sentence is misparsed. A well known example is 'Flying planes is/are dangerous.' Certain syntactic operations, for example assigning subject-verb-object relationships, are probably carried out before meaning is worked out, and it appears that, in English, at any rate, construction of meaning is not carried out before parsing takes place. However, whether parsing takes place first or simultaneously with semantic analysis is a moot point (Rayner and Pollatsek 1989: 263). Rayner and Pollatsek suggest that higher order information drives the whole system. While syntactic parsing and semantic analysis probably proceed simultaneously, however, there may be 'local' structures built up through using syntax, literal meaning of words and heuristics of discourse' (Rayner and Pollatsek 1989: 265) which are then amalgamated to apply to the text as a whole, as illustrated by the model of discourse comprehension presented by Kintsch and van Dijk (1978) and van Dijk and Kintsch (1983), reviewed in Section 1.4.6.1. Sentence structure may help to provide a framework of prediction.

### 1.4.5 Constructing meaning

Rayner and Pollatsek usefully summarise and distinguish the different types of knowledge readers or listeners use to build meaning from a sentence: firstly, the literal meaning of a sentence, for example, identifying agent and recipient, and relationships such as those of time and space; then grammatical rules; then rules of discourse - the way something is expressed rather than the bald what; and finally, pragmatics or real world knowledge (Rayner and Pollatsek 1989: 268). They note that the rules of discourse would not be as strict as the rules of grammar, but may be rules of thumb that have to do with factors such as the limitations of short term memory. By way of example, they cite anaphoric resolution, which they claim to be the simplest task for the reader. An evaluation of research leads them to conclude that grammatical rules are insufficient to decide on an antecedent, and that much of the decision depends on general heuristics or on pragmatic knowledge (Rayner and Pollatsek 1989: 268). My own data suggest that, for some readers of Chinese, anaphoric resolution is a very difficult and uncertain task, and that they depend both on syntactic structures and pragmatic cues to make the decision. This is germane to the discussion of anaphora and is explored more fully in Chapter Four.

### 1.4.6. Propositional representation

Berman (1984) shows that to

*'get at the basic propositional content of a sentence readers must be able to manipulate...constituent structure - what the parts of a sentence are, and how they interrelate hierarchically; structural items - function words and affixes which serve as markers of grammatical relations and of constituent and rhetorical structure; and dependencies - relations expressed between discontinuous elements' (Berman 1984: 140).*

One crucial dependency relationship in text is that of reference and related to it is the reader's cognitive activity of inference.



We can assume three levels of representation of a text: the surface level, at which a reader has a verbatim understanding of the words in their syntactic arrangement; the propositional representation, which encodes the meaning of the text in concepts rather than words; and the mental model constructed by the reader. The hierarchy of propositions of varying degrees of importance in a text creates a story grammar, (eg Thorndyke, 1977) or script (Schank and Abelson, 1977) or frame (Minsky, 1975) for understanding. Some researchers have suggested that if a reader's goal is memorisation, he will focus on syntactic representation, but if it is comprehension he will focus on the propositions and that for the reader, propositions are processing units. The construction of a mental model may depend on the type of text, the type of task and the knowledge of the reader (Stevenson 1993: 109).

Comprehension and memory are closely linked - not only does comprehension pave the way for remembering, but storage of propositions in short term memory is a crucial prerequisite for comprehension. Researchers have found that when recalling stories, subjects remember the gist - that is the hierarchically more important propositions in a text, which are likely to be topicalised or centralised grammatically or discursively (eg Siu 1988, Chen 1990). However, they also tend to remember unnecessary but unusual or surprising detail (Connor 1984) - a notion which can also be related to topicalisation (see Section 4.2.2). On the other hand, in order to understand a text in its entirety, propositions must be held in short term memory while subsequent and dependent propositions are processed.

#### **1.4.6.1 Arguments and propositions within a framework of microstructure and macrostructure**

Kintsch and van Dijk (1978) propose a model of microstructure and macrostructure of text comprehension. They propose two basic units - arguments and propositions; a proposition is a set of arguments having a

predicate as the focus. The microstructure of a text is built up from the linking of propositions but is not simply a list of propositions. The macrostructure is equivalent to the gist of the text, an edited, more profound representation of the text, which the reader ties into his or her schema or long term memory. According to the theory, the reader builds coherence by continually linking propositions that share a common argument. Thus, anaphora and reference are crucial to the reader's comprehension, a notion which will be explored in Chapter Four.

Obviously, linking of propositions and arguments depends a great deal on short term memory capacity. According to van Dijk and Kintsch (1983) short term memory is emptied at the end of each temporal processing cycle, but a few important propositions are retained in order to establish linking for the next set of propositions. The propositions that are retained are likely to be 'high level' propositions, important in the text and to the reader, which are fitted into a pre-existing schema. The reader rejects unimportant material to make way for new, relevant material and adds generalisations and inferences. It is also important to note that while Kintsch and van Dijk see these cycles as coincident with sentences, they do not necessarily see the sentence as an immutable, invariable unit. There is thus scope here for the debatable notion of sentence in Chinese - as discussed later, in Sections 2.5.7.4 and 4.2.2.2, it is a much more fluid phenomenon than the English sentence; in Chinese the clause is perhaps a more reliable processing unit than the sentence, and certainly fits in well with the notions of proposition and microstructure as presented by Kintsch and van Dijk.

#### **1.4.7 The role of schema in reading**

In addition to the lower level processing of graphemes, words, clauses and sentences and the 'middle' level processing of propositions, the background knowledge readers carry with them, as noted above, is crucial to the reading



process. The knowledge a reader already has is used to interpret new information in the text. This knowledge - shaped by and dependent on the reader's experience, and first labelled 'schema' by Bartlett - is closely linked to memory (Bartlett 1932). The reader's schema will help him or her to organise and fill out the information in a text. The more opaque the writing, the more the schema will come into play. For the purposes of reading, schemata may be categorised as formal schemata, or text structure, and content schemata, or background knowledge.

#### **1.4.7.1 Formal schemata**

Perfetti and McCutcheon (1987) argue that comprehension is the link between the meaning of a text and the reader's interpretation. With the briefest of texts - one or two words - reading can stop at the stage of symbols and syntax. Signs such as 'entrance,' 'taxi,' etc. require little elaboration. However, in most reading situations

*'syntactic abilities, coupled with broader scope discourse abilities allow comprehension to proceed by linking together word concepts in such a way as to establish a text base that is a representation of what the text 'says,' even if this is short of a deep representation of the referent world described by the text' (Perfetti and McCutcheon 1987: 108).*

The structure of a text tells the reader what to expect from a text. Layout, punctuation and discourse markers all contribute to the rhetorical organisation and should indicate clearly for the reader temporal, causal, narrative and descriptive sequences. In Chinese texts topic structures, which may be regarded as syntactic or discoursal (see Section 4.2.2), are crucial to text structure as a whole, having a hierarchical mapping (Chen 1986) which should guide the reader in understanding who does what and what is happening.

### **1.4.7.2 Content schemata**

Bartlett (1932) wrote that memory is not 'literally reduplicative and reproductive.... Every time we make it it has its own characteristics' (Bartlett 1932: 204). More recent research has refined and elaborated schema theory; a schema brings together the knowledge built up from first hand experience of the individual and combines with it the vicarious experience of reading, listening and watching. As each new experience or piece of information is gained, it contributes to the construction of the abstract schema and may add to or change knowledge. Understanding of the information supplied in a text interacts with a reader's schemata to bring about interpretation.

It is important to note, however, that miscues at low level may trigger inappropriate schemata. Hudson calls this phenomenon the 'proficiency ceiling:' 'In the process of fitting new input to existing knowledge structures, good reader strategies may cause a break-down in comprehension' (Hudson 1988: 185). This view seems to be 'top down' biased, for if a strategy results in a breakdown in comprehension, the value of the strategy is perhaps questionable. While schemata enable a reader to comprehend the author's purpose, they may also cause a reader to interpret the text in a way entirely different from that intended by the author.

### **1.4.8 Inference**

The application of schemata by a reader may influence mental representation. The more background knowledge readers have the better, as they retrieve information from background knowledge in the long term memory and apply it to propositional representations to help construct their own comprehension of the text.

Where reading is concerned, several kinds of inference play an important part. Here, I will address two main types: bridging inferences and elaborative

inferences. Bridging inferences made about implied information on the basis of explicit information in the text are crucial for comprehension; elaborative inferences, however, are irresistible assumptions made by the reader on the basis of information which is available in his schema, general knowledge or memory, but not set out in the text. Bridging inferences may be problematic where pronouns are absent or ambiguous, as in the case of Chinese text, and particularly so for inexperienced readers, who may fall back on elaborative inferences to achieve a coherent interpretation.

Rayner and Pollatsek point out that much research suggests that readers tend not to draw elaborative inferences and are cautious about drawing any inferences at all, waiting until there is an absolute necessity to do so (Rayner and Pollatsek 1989: 279-281). They suggest that this may be because the costs of a wrong inference outweigh the benefits of a correct one. It may also simply be because the reader wants more information. In any case, it is likely to be skilled readers who exercise this kind of informed patience, while inexperienced readers will be more likely to draw inferences too soon, on the basis of incomplete information.

#### **1.4.8.1 Memory and inference**

In drawing inferences, memory of antecedents is important, and Garrod and Sanford (1977) found that specific concepts in memory are found as antecedents of general concepts more easily than vice versa. This view is corroborated by Duffy's (1986) proposal that causal inference, particularly, is a process of highlighting information and earmarking it for future linking. It is probably easier for a reader to highlight specific information than general information.

It may be worth noting here that some researchers of rhetoric, for example Connor (1984) Kaplan (1966) etc., claim that Western (including English)

writing style tends to progress from the specific to the general, while Eastern (including Chinese) writing style tends to progress from the general to the specific. This is a sweeping generalisation, but there is some evidence that it may be the case with certain types of academic writing at least (Taylor and Chen 1991). If it is true, it has implications for readers of Chinese texts: if a reader is not only faced with a less easily memorable progression from general to specific, and in addition must comprehend a series of null arguments, anaphoric resolution could be more difficult than in a text in a European language. Cui and Chen (1997) found that less skilled native readers of Chinese (those with a lower working memory capacity) were less likely to suppress inappropriate inferences, as described in Section 2.2.

While it is generally agreed that readers tend to remember the gist of a text rather than the detail, readers do often remember some detail, although it is evident from the discussion in the previous section that detail aids memory.

Sachs (1967) tested English reading subjects on their ability to detect changes in surface form, for example, from active to passive voice, in contrasting pairs of sentences such as 'the duchess gave John a painting' and 'John was given a painting by the duchess.' The results showed only chance levels of recognition, leading to the view that in English the surface form of syntax is not retained as part of memory. What is remembered is probably something like a proposition as outlined by Kintsch and van Dijk. Due to the flexible nature of transitivity in some Chinese verbs, reader/translators may assign subject and object roles differently, for example 'X bit Y' or 'Y was bitten by X.' This may be due to readers processing and remembering propositions rather than surface syntax, as Sachs suggests.

It is clear that in comprehending text, the reader loses some detail completely, but it is not a case of all or none. It is likely that some kind of discourse

macrostructure , and as will be discussed in the following sections, mental space structure, is built up from a series of highlighted or important propositions (microstructure) that the reader retains.

### **1.5 Chafe's 'focuses of consciousness and spurts of language'**

Chafe (1980), concerned with the relationship between consciousness and language, suggested the importance of focus: consciousness has a focus and a periphery, maximally activating small amounts of information while making available, but not in focus, larger amounts of information (Chafe 1980: 12). Chafe related this idea directly to his previously explored notion of given and new: 'old information or given information is.....information which the speaker assumes is in the consciousness of the listener at the moment of speaking (Chafe 1980: 10).' Given information is associated with weak stress, and low pitch in speaking, and is the peripheral information, while new information is relatively stressed and emphasised and may be focused upon by the listener.

Chafe suggests that the mechanism of focus, with its varying degrees of activation, is deployed over centres of interest, following paths which are determined by the goals and needs of the individual and his or her already established schemata (Chafe 1980: 12). With regard to reading, Chafe's view is that eye movement studies provide 'independent evidence' for the mechanism of focus of consciousness, associating foveal view with focus, and parafoveal view with peripheral consciousness (Chafe 1980: 13). Chafe cites Just and Carpenter's proposal that fixation in reading reflects what is at the 'top of the stack (Just and Carpenter 1976).'

Where reading is concerned, fixation, and thereby focus, may be associated with content words, that is, relatively long, complex or dense orthographic items of relatively low frequency, while saccades, and thereby peripheral

consciousness, may be associated with function words, that is simpler, less dense orthographic items of high frequency. This notion is discussed more fully with regard to empty words and zero pronouns in Chinese in Section 2.5.7 and Section 3.5.

Chafe equates the focus of consciousness with 'spurts of language' - that is, what he calls an 'idea unit' which corresponds to a clause or a part of a clause. According to Chafe, the 'idea unit' consists, syntactically, of 'one verb with whatever accompanying noun phrases are associated with it (Chafe 1980: 14).' Noting the limited capacity of human consciousness, Chafe suggests that the idea unit which equates to a focus of consciousness may be too limited to be useful, such that several idea units are needed to scan information that is being attended to, and that these several idea units may be called a centre of interest. The limitations of short term or working memory, and its relationship to reading processing are discussed in Section 2.2. In terms of language, the centre of interest may be equivalent to a sentence, closed by a final (probably falling) contour, and containing several clauses or idea units (Chafe 1980: 26). This notion may be very relevant for Chinese clauses and sentences. Later writers such as Fox (1987) and Liang (1996) have developed the centre of interest as a key component of a hierarchical rhetorical structure.

## **1.6 Mental spaces**

The mental space outlined by Fauconnier (1994) equates in role to a mental model but Fauconnier rejects the notion of proposition. This may be relevant to Chinese, as the phenomenon of the null argument creates predicates that have no argument, thus questioning the whole notion of proposition. Fauconnier's initial motivation was to solve difficulties in finding 'coherent, logical representations of pronominalisation phenomena' (Fauconnier 1994: 162).



Mental spaces are not mental models of the world, but mental models of discourse. According to Fauconnier, highly abstract mental creations support talk and reasoning; grammar guides but does not define these mental creations (Fauconnier 1994: xxiii). He claims that as a 'semantically heterogeneous and highly unspecified form' the sentence does not carry propositions. Rather, it gives simultaneous information about different aspects of discourse-building and meaning construction. For Fauconnier language simply reflects conceptual connections and linguistic items do not bear meaning, but act as steps in constructing meaning (Fauconnier 1994: xxiii). Fauconnier's mental spaces seem to be not incompatible with Chafe's foci of consciousness and centres of interest (see 1.3. above).

Fauconnier argues that the schemata, conceptual connections, metaphors and so on, of the real world, are much richer than the configuration of language (Fauconnier 1994: xxvii). He maintains that mental spaces are different from domains and possible worlds in that mental spaces are only partially specified, indeed 'extra-ordinarily underspecified,' whereas domains and possible worlds are fully specified (for example, the world of the viewer and the world of the film viewed, as described in the pear stories (Chafe 1980: 46)). We use mental spaces to think and talk about domains. But there is not a one to one correspondence between mental spaces and linguistic forms - events in the real world, the social situations of discourse, previous discourse and its linguistic forms and cross space mapping all contribute to mental space construction: 'The power of grammar is to call up suitable generic frames that will serve in context to manipulate much more specific ones' (Fauconnier 1994: xxxvi).

Fauconnier believes that interlocutors pursue the 'maze of spaces' by means of tenses, moods, space builders such as time and space adverbs and anaphors. How could his theory, then, possibly be applied to Chinese, in

which, arguably, there are no morphological tense markers, only aspect markers, there is no subjunctive mood, sometimes no distinction between active and passive voice, transitivity is flexible, space builders such as conjunctions are often optional and the zero anaphor is the unmarked form in discourse?

Perhaps it is the very underspecification of the mental space that is appealing and applicable to Chinese. The absence of space builders and connectors (and even verb operators, as will be seen in Section 4.2.1.2 - Chinese not only has a tendency to null arguments, but in some cases verbless sentences) could provide more open access to and between mental spaces. Fauconnier does note in passing that while space tracking by means of grammatical devices may be universal across languages, indications provided are language specific (Fauconnier 1994: xli). The 'relatively simple grammatical structures' that Fauconnier suggests give underdetermined instructions for space construction, in Chinese are often not just simple, but minimal. Many of Fauconnier's English examples reveal a high degree of ellipsis and of 'sloppy identity,' as follows:

1. 1. Plato is on the top shelf. It is bound in leather. You'll find that he is a very interesting author.
2. Mary did not manage to catch a fish, otherwise she would have cooked it. She might have given it to her children.

(Fauconnier 1994)

Fauconnier notes that a major difficulty is the treatment of variables, especially as anaphors bind an antecedent even if they are outside the logical scope of the operator, and 'a pronoun with an antecedent in one space can freely identify with its counterpart in another connected space' (Fauconnier 1994: 35). In both the examples given above, the perfect solution lies in



Chinese - the zero pronoun. The verbs provide all the information necessary to access the requisite mental spaces. So it could be said that in Chinese the zero pronoun, far from being a hindrance, may be a help and may provide greater coherence. For example:

2.

|  |          |            |        |         |
|--|----------|------------|--------|---------|
| 这 颗  | 树 叶子     | 没有 了       | 那颗     | 还有      |
| zhei ke  | shu yezi | mei you le | nei ke | hai you |
| this (measure) tree leaves not have (asp.) that (measure) still have |          |            |        |         |
| This tree has lost its leaves, that one still has them.              |          |            |        |         |

In certain cases, then, a zero anaphor or null argument may provide links between mental spaces, where the use of an overt pronoun could set up barriers between spaces. This shows how minimal the information needed to construct a mental space can be. Fauconnier notes 'Many linguistic reference phenomena do not depend on which particular connectors are involved, but rather on general properties of open connectors and the trigger-target properties they set up' (Fauconnier 1994: 19). He further notes that where a pronoun is ambiguous, the availability of two interpretations is a consequence of the discourse processing involved, rather than a structural, semantic or logical ambiguity; the double processing possibility will appear if there is more than one space.

Fauconnier notes that many verbs either establish relations between spaces or set up new spaces. It may be that in Chinese the verb may carry greater responsibility as the trans-spatial operator, owing to the elliptical nature of Chinese in other respects. Could it be that verbs trigger mental spaces directly, while the very absence of connectors and space builders allows for greater accessibility to and connection between spaces? This notion is discussed more fully in a Chinese context, in relation to the data, in Chapters Four, Five and Six.

## 1.7 Summary and conclusion

The debate about how reading happens, and how it should be taught continues but may be aided by empirical research. Huey's (1908) study is still instructive in that it addresses many important concerns and adumbrates later experimental work.

While the notion of 'bottom up' processing may be useful for recognising or working out words in alphabetic writing systems, Chinese until now has mainly been taught by memorisation of single characters. Initial teaching may differ, but with both types of orthography perception of print becomes an automatised process and as such is not a central concern of this study.

A more complete picture of reading is provided by models of reading which take account of higher level processes such as that first proposed by Goodman (1967, 1975) and subsequently labelled a 'top down' model, and later interactive models such as those of Rumelhart (1977) and Grabe (1988).

Experimental studies have informed reading theory at all levels including and beyond that of word or character recognition. For example, eye movement studies have shown that readers of English tend to skip function words and fixate content words (Rayner and Pollatsek 1989) and this phenomenon may also be applicable to reading Chinese. Eye movements have some bearing on syntactic processing but even so, there is still disagreement as to whether syntactic processing precedes semantic processing or vice versa (Rayner and Pollatsek 1989). Sentence structure seems to aid in prediction as local structures build up into textual structure (Kintsch and van Dijk 1978, van Dijk and Kintsch 1983). In addition to syntactic rules discourse rules also play a part in reading, for example, in short term memory, and in tracking given and new information (Chafe 1980). For the purposes of understanding a reader probably focuses on propositions to construct a mental model of some kind.

This may consist of a macrostructure (or gist) and a microstructure (sequence of propositions) (van Dijk and Kintsch 1983).

Readers probably reject unimportant material and retain important material, building it into their pre-existing schemata. In reading English, processing cycles may coincide with sentences. Text structure, or formal schema, tells the reader what to expect from a text, aiding prediction (Perfetti and McCutcheon (1987). Background general knowledge, or content schema (Bartlett 1932), helps a reader fill out the information in a text. Both bridging inferences and elaborative inferences play a part in reading, and memory of antecedents aids both (Garrod and Sanford 1977, Duffy 1986).

Chafe (1980) outlined a model of centres of interest, which, like van Dijk and Kintsch's microstructure, equates to the syntactic units of a language. Chafe's idea units correspond to clauses, and several idea units, possibly a sentence, correspond to a centre of interest. Old, or given information is assumed by the speaker or writer to be in the consciousness of the listener or reader. Chafe associates focus of consciousness with foveal view in reading and peripheral consciousness with parafoveal view.

Fauconnier (1994) rejects the notion of propositions, arguing that linguistic items do not bear meaning but act as steps in constructing meaning (Fauconnier 1994). A key notion of mental space is that it is underspecified and may be very suitable as a model for the comprehension of elliptical elements such as the null argument or zero pronoun in Chinese.

### **1.7.1 A note on the usefulness of experimental studies of reading**

While there is no doubt that empirical studies of reading are essential in supporting theories, are they as authoritative as experimenters claim? Rayner

and Pollatsek, for example, go so far as to suggest that reading inside the laboratory is virtually the same as reading outside the laboratory (Rayner and Pollatsek 1989: 23) but a glance at the methodology of their experiments and those of other researchers may give one an opposite impression. While they admit that eye movement studies are not the only way to study reading, they insist on the usefulness of eye movement studies. They claim that 'where readers look and how long they look there provides valuable information about the mental processes associated with understanding a given word or set of words' (Rayner and Pollatsek 1989: 23). However, normal texts do not consist of rows of crosses or boxes interspersed with occasional words or characters, as 'moving windows' experiments do. Reading from a computer screen is definitely now normal reading, but although many young people read almost nothing but computer screens, both for work and leisure, when they do so in daily life, they read whole meaningful texts, not disembodied or manipulated texts. My own feeling is that eye movement studies may serve to corroborate other evidence of the processes of reading, but should not be relied upon as sole evidence.

Reductionist experiments of all kinds have provided valuable information about parts of the process of reading. Tasks designed by experimental psychologists tend to take distinct activities or sub-skills such as word recognition or parsing and isolate them from the rest of the reading process, and how reading is studied depends very much on the part of the process in which the particular researcher is interested. The value of the research is to mix and match, wherever possible, the various parts into a feasible whole.

## **Chapter 2: Chinese reading: traditional approaches, characteristic features of Chinese text and recent Chinese psycholinguistic studies of reading**

### **2.0 Aim and introduction**

The aim of this chapter is to explore aspects of reading which are particularly relevant to the reading of Chinese, and some aspects which may be exclusive to the reading of Chinese. As the aim of the thesis is to investigate reading of the null argument or zero pronoun, most lower level processes are not discussed, except where they affect higher level processes such as parsing. So, problems such as the debate about phonological and visual access of Chinese characters will not be covered. It is assumed, however, on the basis of the body of research on the subject (Tzeng et al. 1981, 1986, 1995, de Francis 1989, Zhang and Simon 1985 etc.), that readers access Chinese characters both phonologically and visually, and that miscues which affect other levels of processing may arise in both cases.

A very brief description of Chinese language and writing is given as background to the reading studies reviewed in this chapter, and also to the discussion of translation in Chapter Three and the null argument in Chapter Four.

In the twentieth century, Chinese psycholinguists began to look at reading processes in a descriptive, empirical way. Although for the classroom teacher, the key issue in reading is still the recognition of characters, of which thousands must be learnt in order to achieve literacy<sup>4</sup>, reading processes must also include cognition at syntactic and semantic levels. The research reviewed includes studies of reading achievement in schools in Hong Kong, Singapore

---

<sup>4</sup> Basic literacy requires 1,500 characters. Around 6,000 are needed, for example, to do a higher education course.

and Taiwan, and experimental research on the reading of adults and children by means of eye movement studies, recall and comprehension studies. Some of these studies investigate the importance of animacy, word order, the meaning of the verb, text structure and thematic information in constructing meaning for the whole text.

**2.1. About the Chinese language**

Before embarking on an investigation of reading Chinese text, it is necessary to outline certain important characteristics of the language and the orthography which are relevant to any study of the reading process.

**2.1.1. Chinese ‘grammar’**

One of the most distinctive characteristics of Chinese is that it has virtually no morphological inflection. Chinese words do not ‘change’ in the way that words in European languages do - a Chinese word cannot be given additional or changed meaning by the application of a prefix, a suffix or an infix. Verbs do not inflect for person, number or tense, though there are aspect markers in the form of additional syllables. eg

|    |                                   |                 |                      |
|----|-----------------------------------|-----------------|----------------------|
| 3. | 他吃                                | 他们吃             | 吃过                   |
|    | ta chi                            | tamen chi       | chiguo               |
|    | he eats/ate                       | they eat/ate    | I/you/he/she/we/they |
|    | is eating/ etc.                   | are eating etc. | have eaten           |
|    | 吃了                                |                 | 吃著                   |
|    | chile                             |                 | chizhe               |
|    | I/you/he/she/we/they              |                 | I/you/he/she/we/they |
|    | have eaten, ate, are about to eat |                 | are/were eating      |

Thus time is always expressed in terms of time adverbials, possibly with an aspect particle. The passive voice is rare in Chinese (though on the increase) and is formed by the addition of a co-verb:

4.     他打我                      我被他打  
        Ta da wo                 Wo bei ta da  
        he hit me                I    by him hit  
        He hit me                I was hit by him.

Chinese nouns do not inflect for gender, number, or case:

5a.

|                   |                      |                    |
|-------------------|----------------------|--------------------|
| 学 生               | 给 老师                 | 买 书                |
| xuesheng          | gei laoshi           | mai shu            |
| (a/the) student/s | for (a/the)teacher/s | buy (a/some)book/s |

5b.

|                  |                       |                     |
|------------------|-----------------------|---------------------|
| 老师               | 给 学生                  | 买 书                 |
| laoshi           | gei xuesheng          | mai shu             |
| (a/the)teacher/s | for (a/the) student/s | buy (a/some) book/s |

5c.

|        |             |                             |
|--------|-------------|-----------------------------|
|        | 书           | 卖 完了                        |
|        | shu         | maiwanle                    |
| (when) | (the)book/s | (were/are/will be) sold out |

Chinese adjectives do not inflect for gender, number or case:

6a.

那位 小姐 很 好看  
 neiwei xiaojie hen haokan  
 that girl very good-looking  
 That girl is very good looking.

6b.

那位 先生 很 好看  
 neiwei xiansheng hen haokan  
 that gentleman very good-looking  
 That man is very good looking.

There are no definite or indefinite articles as such - if definiteness needs to be expressed, it takes the form of demonstratives such as 'this,' 'that,' 'these,' 'those' etc.

Chinese can be extremely succinct , especially in its written form, and more especially in literary writing. Much of the information which is carried by morphological inflection in European languages may simply not be there in Chinese.

Most Chinese linguists regard Chinese as having a topic prominent structure rather than a subject predicate structure, but there is still scope for the subject object relationship, as described in Chapter Four.

Pronouns and connectors are not always used. Some linguists regard the zero pronoun or zero argument as the normal mode of anaphora in Chinese. The zero pronoun, or null argument plays an important part in the continuity of the topic chain. This is discussed in greater detail in Chapter Four.

### 2.1.2 Chinese language and writing

Modern Chinese writing relates to the spoken language in such a way that there is one character per syllable and one syllable per morpheme, and in many cases, this character/syllable/morpheme unit is also equivalent to a word. No changes can be made to a character - if an aspect suffix is required



it is provided by an additional character, as seen in the examples above. Polysyllabic words are made by adding further morphemes/characters:

7.

|                              |            |     |
|------------------------------|------------|-----|
| vehicle                      | che        | 车   |
| train (fire vehicle)         | huoche     | 火车  |
| station (fire vehicle stand) | huochezhan | 火车站 |

Chinese characters consist of a radical and a phonetic. A radical may be a character/word in its own right. Radicals sometimes give clues to meaning, and phonetics sometimes give clues to sound, but neither can be relied upon on in all cases:

8.

|            |   |              |   |       |
|------------|---|--------------|---|-------|
| 心(radical) | > | 相 (phonetic) | > | 想     |
| xin        |   | xiang        |   | xiang |
| heart      |   | image        |   | think |

Learning to read Chinese is usually achieved by memorisation of whole characters. Every character has a pronunciation rooted in the spoken language, so reading is not a purely visual activity.

Chinese may be written left to right horizontally, as in the PRC and Singapore, or right to left vertically as in Taiwan, and in both cases characters are equally spaced - where there are polysyllabic words there is no inter-word spacing.

**2.2 Chinese language and short term memory**

Short term memory, or working memory span affects all language processing. In reading , where the text is present for the reader to carry out retrogressive monitoring, it is still essential to retain previous information which contributes semantically and syntactically to the ongoing process.

Zhang and Simon (1985) compared short term memory spans for English and Chinese language materials, basing their study on measurement in terms of chunks as proposed by Miller (1956) . For Chinese language materials they found a non-phonological component of short term memory with a capacity of about three chunks, corresponding to Baddeley's 'visuo-spatial scratch pad' (Baddeley 1983).

Acknowledging the difficulty of finding suitable equivalents for measurement in Chinese and English, Zhang and Simon aimed to obtain the following:

1. The STM span for three basic units in Chinese writing - radicals, characters and words;
2. The effects on STM span of pronouncability for visual stimuli;
3. Measurement of acoustical compared with non-acoustical STM;
4. The effect on STM of the number of syllables in a stimulus item.

(Zhang and Simon 1985: 196)

They predicted a similar STM capacity for radicals, characters and words. They found that for unnamed radicals and homophonic characters, readers' recall was about three items. For non-homophonic characters, they found an STM span of seven items (Zhang and Simon 1985: 196). A high rate of homophonic intrusions and a low measured span of non-acoustic STM led Zhang and Simon to challenge the popular notion that Chinese readers encode characters 'directly' from visual to semantic form, without an intermediary acoustic encoding (Zhang and Simon 1985: 196). Written Chinese is undoubtedly linguistic in nature and oral language is an essential intermediary for readers of both alphabetic scripts and Chinese script.

Zhang and Simon also found that familiarity was an important factor in recall: acoustically retained material is likely to be encoded in terms of the most frequent grapheme in that homophone group (Zhang and Simon 1985: 198), a

notion which is compatible with Glushko's theory of families in word recognition (Glushko 1979).

Just and Carpenter suggest that reading span reflects working memory capacity, and that a reader's memory capacity will influence reading processes. Readers with a good working memory are likely to retain more information (Just and Carpenter 1992). In support of this, Cui and Chen tested groups of native Chinese readers for working memory and then administered reading tests. They found that those who had a higher working memory capacity named characters faster and drew fewer predictive inferences. Their lower working memory group named characters more slowly and were less able to suppress inappropriate inferences (Cui and Chen 1997).

### **2.3 Chinese language and reading disability**

A common misunderstanding in the West is that dyslexia does not exist among Chinese and Japanese reading subjects. However, a number of studies have highlighted particular difficulties associated with reading Chinese, showing that reading disability does exist, and information about it may shed light on reading processing in general.

Stevenson, Stigler, Lucker, Lee, Hsu and Kitamura (1982) carried out a comparison of reading disability among children in three comparable cities, Taipei in Taiwan, Sendai in Japan and Minneapolis in the USA. It was found that the Chinese reading subjects showed more general disability in learning than the American and Japanese subjects, who showed more specific reading problems. Where children failed the fifth grade test and had to take the fourth grade test, Chinese readers were more likely to pass the vocabulary test and fail the comprehension test, whereas American and Japanese readers were more likely to fail the vocabulary test and pass the comprehension test

(Stevenson, Stigler, Lucker, Lee, Hsu and Kitamura 1982). This could be explained by the dominance of memorisation of Chinese characters, or by the fact that English and Japanese texts provide much more linguistic information by means of inflectional morphology.

Stevenson (1984) and Stevenson et al. (1985) studied first grade children in the same three cities, finding that while the Chinese speaking children did far better in maths, and read better than Americans or Japanese, American and Japanese children did better in cognitive tests (Stevenson 1984, 1985), a result which is attributed by Hoosain (1991: 76) to the fact that Chinese children work much harder that is, they memorise.

Dissatisfied with Stevenson's results on the grounds that only standardised tests were used, Su, Soong, and Hsu (1984) carried out a smaller study using the written protocols of twenty children. These subjects showed five common error patterns: 1. Confusion of words with a similar (visual) configuration; 2. Phonological confusion of words with similar pronunciation; 3. Confusion of the position of the radical within a character; 4. Character order confusions or reversal of characters within a common vocabulary word; 5. Mirror writing.

The first four of these error types I have found in my data among competent readers who showed excellent understanding of texts in other respects (Waggot 1997), and the fifth error type is often associated with phenomena such as cross handedness (Chinese children have traditionally been forced to use their right hands for writing) and certain types of aphasia (Hoosain 1991: 147). All twenty of Su's subjects had a higher performance IQ and perceptual organisation than verbal IQ. Su et al. concluded that one possible cause of reading disability might be due to brain dysfunction (Su et al. 1984). This type of study appears to test only recognition and writing of individual

characters, rather than comprehension of linguistic structures in connected text.

A further follow up study was carried out by Lee, Wee and Wong (1987) on Chinese background Singaporean children who learn to read and write Chinese and English simultaneously. They found that verbal memory in Chinese and conceptual ability discriminated good and poor readers in Chinese in both vocabulary and comprehension. Verbal memory was found to be important in reading both languages and had a low positive correlation between the two languages, which they attributed to auditory comprehension and recall of facts heard verbally. This is in line with other research which suggests that working memory span contributes significantly to English reading ability (Baddeley, Logie and Ellis 1988).

However, contrary to findings which suggest that reading is basically a linguistic skill in which visuo-spatial skills are less important (Rutter and Yule 1976; Vellutino 1987), Lee et al. found that scores on a conceptual ability test which required perceptual reasoning were positively correlated with reading ability (Lee, Wee and Wong 1987). Tzeng and Hung (1988) explain this by saying that Chinese logographs may be read by a process of generating phonological codes by analogy. Readers can make guesses about the pronunciation of newly encountered characters which share the same phonetic component, a procedure different from grapheme-phoneme conversion rules outlined by Coltheart (1980), but similar to Glushko's (1979) activation model.

What is not clear about the bilingual studies such as those of Lee et al. is the extent to which the two languages influence each other. It is evident from Hsia and Leong's observations (reviewed briefly in Section 4.1) that English has a profound influence on Chinese in Hong Kong, and the same may be

true of Singapore, where English medium education has until now been more widespread.

## **2.4 Levels or stages in interactive reading as applied to Chinese**

As described in Chapter One, reading consists of parallel and interdependent processes at various levels from word or character recognition through to application of background knowledge. While the issues of character recognition and lexical access are outwith the scope of this thesis, where my data suggest that miscues at these levels appear to have affected higher level processes, the miscues and their effects are dealt with on a case by case basis in Chapter Five.

### **2.4.1 Some perceptual aspects of reading Chinese text - density, complexity, regularity and spacing**

Quite apart from the way in which Chinese characters are composed and read, Chinese is noticeably different from alphabetic scripts in the way it is arranged on the page. Woodworth (1938) went so far as to suggest that the 'box-like shapes might be better adapted to exploit the small disc fovea in the retina than alphabetic words with their rectangular outlines' (Woodworth 1938, cited in Sun and Feng 1999).

Spacing and punctuation guide the reader to an awareness of word and sentence boundaries. Most European languages are spaced between words (as conventionally perceived) and have a capital letter at the beginning of proper nouns and the beginning of sentences (as conventionally perceived) and a full stop at the end. German gives every noun an upper case initial. Some African languages space between phrases. Chinese and Japanese space between syllables (with both kana and kanji in the case of Japanese) and have no upper case to indicate either sentence initial or proper nouns.

Woodworth’s view (above) is reflected, but not supported, in the work of Chen (1996). Chen points out how heavily readers rely on visuo-spatial information in a text, for example, the height of letters and the length of words. Citing Western research he explains how spatial layout factors affect performance in reading alphabetic scripts. Just, Carpenter and Woolley (1982) found that, in alphabetic scripts, the longer the word, the longer the reading time. In addition, saccades to the right of fixation tended to be longer when a long word is encountered (Yang and McConkie 1999). Chen demonstrates how reorganisation of spacing and upper and lower case letters, as in :

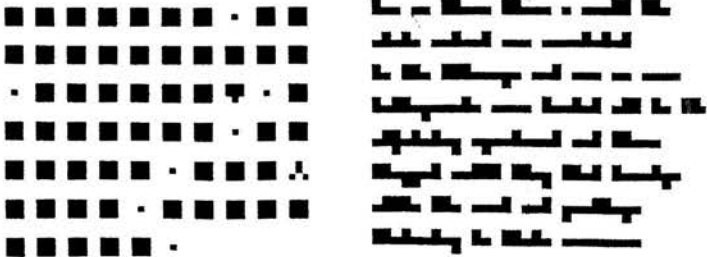
‘tHiSsEnTeNcEiSdIfFiCuLtToReAd’

may disturb a reader’s pattern of processing and comprehension (Chen 1996: 44).

Chen illustrates the visual spatial differences between Chinese and English by means of a quantization procedure which removes the fine detail from print leaving ‘global’ spatial information.

9.

|   |   |
|---|---|
| 每月總有兩三次，全校<br>學生列隊在圖書館集合<br>，總有幾個壞孩子，會<br>給推進隔壁房間，鞭打<br>得皮破血流，其餘的人<br>坐在那裡，哆哆嗦嗦聽他<br>們呼號尖叫。 | Two or three times a month the<br>whole school was marshaled<br>in the library, and one or more<br>delinquents were hauled off to the<br>adjoining apartment and there<br>flogged until they bled freely,<br>while the rest sat quaking,<br>listening to their screams. |
|---|---|



(Chen 1996)

His message is that Chinese characters can be seen as strings of square blocks with some exceptions. However, the quantization of the exceptions (those



with few strokes) is not consistent and my feeling is that the 'box shaped' elements are not as constant as he claims. If readers are influenced by any global spatial information in a character it is likely to be density. What is certain in Chen's argument, however, is that the overall visual pattern of Chinese characters is far more regular than that of alphabetic script, in Chen's example, English. While English words vary greatly in length and extend irregularly below and above the line, the Chinese characters are spaced equally, except where a punctuation mark is inserted. The Chinese text shown in Chen's quantization illustration is elsewhere used for experimental purposes and is presented unusually, in that the characters are widely spaced. In authentic printed texts, spacing is often minimal, as in the text below:

10.

列車有些顛簸，所以他看上去搖搖晃晃的。他手中提著一只輕飄飄的白色牙缸，見我要進洗漱間，就用空牙缸指了指水龍頭，然後竭力地搖頭，我便明白水已經用光了。他依然一絲不苟地刷牙，唇角溢出的牙膏沫越聚越多，一股清香味撩人地漂浮著，仿佛他使用的牙膏是用青草榨成的。只有心境平和、生性寬厚的男人在處理此類尷尬事時才等閑視之。我不知道他最後如何處理那些牙膏沫，便回到車廂把喝剩的半瓶礦泉水拿過去。他主動把牙缸伸過來，很自然地讓我把礦泉水咕咕地傾倒進去，然後俯身在洗臉池前，飛快地把口腔沖洗乾淨。

The noticeable spaces come at the ends of sentences, where full stops occur, and in some cases the ends of clauses, where commas or semi-colons occur. A punctuation mark in a Chinese text is allotted the same space as a character. In effect Chinese characters in a printed text are set out continuously, though not contiguously.<sup>5</sup> While some low density characters may produce a partial space, as shown in Chen's quantization, true spaces only occur at syntactic breaks. The issue of the 'gap' in Chinese text will be considered more fully later, but it may be the case that spaces in text provide different information

<sup>5</sup> Like handwriting in most languages, Chinese handwriting, 'grass script' in particular, lacks the



for readers of Chinese characters and readers of alphabetic scripts. Moreover, it may be the case that for inexperienced readers, regular spacing may mislead in parsing.

#### **2.4.2 The psychological reality of the Chinese word - a word or a constituent?**

The character is recognised as the sociological word - 'that type of unit , intermediate in size between a phoneme and a sentence , which the general non-linguistic public is conscious of.....'(Chao 1968: 136). Each character is learned individually and holistically. As discussed above, a character may represent a word, and many words in Chinese are monosyllables and are written as a single character. (One character = one syllable = one morpheme = whole or part word.) However, there is a Chinese word for 'word' and the Chinese language is considered to include words of more than one syllable or character. Experimental studies have shown that, while Chinese reading subjects are aware of the notion of word, they are not always sure where a word begins and ends. Verbs are particularly opaque in this respect, for example, while YAO QU 要去 (will go) would probably be regarded as two words, QU GUO 去过 (have gone) might be regarded as one word, or as a full word followed by an empty word (see discussion of full and empty words in Section 2.5.7.1).

Chen (1996) carried out experiments designed to determine whether Chinese has a word superiority effect, as found to be the case with English, where, for example, the letter *k* is more easily seen as part of the word *work* than when standing alone (Huey 1908: 72). Chen claims that the Chinese word plays a distinctive role in Chinese reading. He found that when subjects were asked to retain (and respond to) the occurrence of a given character in a sentence, response latencies were longer when the target character was embedded in a

non-word rather than a word context. Cheng (1981), using tachistoscopic presentation, obtained similar results. Chen concludes that this word superiority effect indicates that Chinese readers do have mental representations corresponding to words, as well as to characters (Chen 1996: 57).

Mattingly and Hsiao (1997), however, claim that word superiority is in fact constituent superiority and 'inheres in any orthographic unit or grouping that could stand for a grammatical constituent of an utterance; a morpheme, a word or phrase' (Mattingly and Hsiao 1997: 215). This is a very appealing argument, given the uncertainty about 'words' in Chinese. However, Mattingly and Hsiao's experiment seems somewhat unreliable as very few of their subjects carried out the task according to instructions; when asked to select radicals, in many cases they selected whole characters which the experimenters then computed as correct responses, along with the radicals. By way of explanation, in other situations, it has been found that Chinese subjects 'play safe' by providing more information than is necessary in an answer. For example, it has been found that Chinese readers, when answering 'short answer' questions in reading comprehension tests, tend to copy out chunks from the text which exceed, but include, the answer (Lewis, personal communication). In research on Chinese speakers' reactions to lie tests, it has been found that subjects may prefer to give a wrong answer than to admit to not knowing (Yang and Chiu 1988). These two observations could indicate that Mattingly and Hsiao's subjects were uncertain and chose to err on the safe side.

Hoosain (1991), points out that

*'the character could have an ambivalent role as the unit of perception in Chinese. Whereas it could fuse its identity into a whole within two-character words, there are times when it retains its individual salience as a result of its status as a sensory unit' (Hoosain 1991: 31, 32).*

Chen (1996) remarks that because words in a Chinese text are not explicitly marked as such, while they may not serve as encoding units, like alphabetic words, they may play an important role in comprehension and semantic analysis. Liu, Yeh, Wang and Chang (1974), however, found no advantage for readers when they were provided with gaps between words in a Chinese text, though this may have been due to life long habit, and no opportunity to get used to a novel format (Liu, Yeh, Wang and Chang 1974).

#### 2.4.2.1 *chengyu* as words

The meaning of a word is generally related to the meaning of its component characters, except in the case of idioms (*chengyu* 成语) (Chen 1996). Hung, Tzeng and Chen (1996) found that different types of Chinese words may activate different recognition processes. They found that multi-character words (idioms or *chengyu*), in which there is not necessarily a semantic connection between the word itself and its constituents, are likely to be processed as a single psychological unit. For example, the following *chengyu* which translates as 'indistinguishable':

11.

|                     |       |     |          |
|---------------------|-------|-----|----------|
| 玉                   | 石     | 不   | 分        |
| yu                  | shi   | bu  | fen      |
| jade                | stone | not | separate |
| 'indistinguishable' |       |     |          |

*Chengyu* are basically metaphors and are sometimes referred to in English as set phrases or four-character phrases, as this is their most common format. The fact is that these multi-character 'words,' often containing a whole sentence structure, need to be memorised, either by frequent use, or by rote, as in a classroom situation. Inexperienced speaker/readers of the language may not necessarily recognise them as single units, but may attempt to read them character by character, with disastrous results - an example of opacity at

word or phrase level rather than at character level. A *chengyu* may function as any part of speech - noun, verb, adjective, adverb - without changing in form. A number of these problematic items occur in the texts used in my study, in some cases as predicates of null subjects, and will be discussed in Chapter Five.

The unreliability of either the word or the character as a stable processing unit in Chinese influences Chen (1996) in his view that Chinese readers may use a more diffuse reading strategy than readers of alphabetic scripts, as demonstrated in the violation experiments, discussed in Section 2.6.3.1.

Readers of English and other European languages not only have help in the demarcation of words, but have additional help in the information provided by inflectional endings, derivational prefixes and suffixes and so on. Readers of Chinese, however, have only a very vague mental representation of a 'word' and are not shown by their writing system where it begins or ends, or how it relates to other words in the text.

## **2.5 Eye movement studies of Chinese reading**

As described above, eye movement studies have been successfully used to investigate processing of units larger than words in languages using alphabetic scripts and also in Chinese.

### **2.5.1 A note on scanning direction in Chinese reading**

Before considering eye movement it is necessary to deal briefly with the question of horizontal versus vertical reading. Individual Chinese characters are written in a defined stroke order from left to right and top to bottom, except when they are composed symmetrically. Text consisting of Chinese characters, however, may be written and read either vertically from right to left or horizontally from left to right. Very short texts, such as headlines, shop signs, and so on, may also be written and read horizontally right to left. It has

been found that Chinese reading subjects' scanning preference depends on reading experience. Readers who have been brought up to read vertically, as in Taiwan and Hong Kong, read vertical arrangements faster and more efficiently than horizontal arrangements, and readers who have been brought up to read horizontally, as in mainland China and Singapore, read horizontal arrangements faster and more efficiently than vertical arrangements (Chen and Chen 1988). Bilingual readers who are accustomed to reading Chinese vertically have no problem becoming accustomed to reading English horizontally (Chen 1996: 47). It can therefore be assumed that any difference in directional scanning does not have psychological implications for comprehension of text.

### **2.5.2 Saccades and fixation and their relationship to reading comprehension**

Some researchers have claimed that readers do not always use a saccadic pattern of eye movement when reading Chinese text. Shen (1927) and Stern (1978) both claimed that Chinese readers use the pursuit pattern of eye movement more often associated with tracking moving objects. Stern's subjects, however, did not all show this pattern consistently.

It has been shown that when displaying a pattern of saccadic movements, Chinese readers tend to make smaller and more regular saccades than readers of alphabetic scripts, as might be expected given the greater visuo-spatial regularity of Chinese script, as described above (Chen 1996: 49). There may be additional reasons for the smoothness and regularity of eye movement pattern in reading Chinese. Chen points out that in visuo-spatial terms a single Chinese character may be more complex (that is, consist of more lines or strokes) than a single letter, while a single alphabetic word may be much longer than a single Chinese character. Moreover, owing to the absence of inflectional information, Chinese characters bear a greater informational

density. Chen points out that in alphabetic scripts, the great variation in the visual form of words - length and shape - provides parafoveal information which guides the reader to appropriate saccades, in effect giving the reader a parafoveal preview of the forthcoming word. The extreme regularity of Chinese script, however, provides relatively little parafoveal information. It is possible that the small regular saccades of a Chinese reader's eye movement 'glide into a pattern of smooth pursuit like movements' (Chen 1996: 49).

### **2.5.3 Moving window presentation**

Moving window presentation has been used as a cheaper, more convenient substitute for conventional eye movement equipment. Sun and Feng (1999), used a sliding presentation on computer screen which allowed subjects to select their own speed of reading. In their analysis of three groups of subjects, primary school pupils, high school pupils and graduates, they found a similarity of eye movement pattern in reading English and Chinese. Their findings suggested that fixation and recognition are controlled not by the visuo-spatial characteristics of the script, but by the linguistic information of the language. Their conclusion was that any limitation or difficulty in reading was 'related to high level cognition rather than the peripheral nervous system' (Sun and Feng 1999: 201).

### **2.5.4 Preferred landing position**

Yang and McConkie's (1999) investigations of the patterns of eye movement in reading Chinese covered a greater range of variables, moving away from the simple factors of speed and movement to frequency distributions of the patterns. They also investigated the possibility that it is the word in Chinese, rather than the character, that guides the reader's eye movement. Finally, they examined the effect on eye behaviour of character complexity, word complexity and word frequency (Yang and McConkie 1999).



The preferred landing position of the eyes when reading English has been found to be slightly to the left of centre of a word. If a similar phenomenon existed in reading Chinese, it would necessitate the location of words in peripheral vision prior to saccade and landing (Yang and McConkie 1999: 210). Yang and McConkie point out that it is likely that high frequency, less complex characters (ie. empty, or function words) are likely to be fixated for shorter periods than low frequency, more complex characters (ie. full, or content words). They suggest that luminance patterns affect landing positions, and a Chinese reader is less likely to skip high density, complex characters. This suggestion would support the notion that Chinese readers tend to skip empty words, as discussed later in Section 2.5.7.1.

Yang and McConkie tested the eye movement patterns of 13 Chinese reading subjects reading words of high and low frequency with varying patterns of high and low frequency characters and high and low complexity characters (Yang and McConkie 1999: 210). They found that among their excellent readers, saccade length differed substantially, ranging from 1 - 14 units, (where one unit is equal to half the width of one character) but that the average saccade length was 2 characters. They also found that there was no particular preferred landing position and that readers' gaze landed on spaces as frequently as on characters. These findings led them to the conclusion that there is no evidence that the reader of Chinese is guided by the visuo-spatial characteristics of either words or characters (Yang and McConkie 1999: 215).

They do not specify what type of space their readers' eyes landed on, that is, whether it was simply the typographical spaces between characters, or whether it was the syntactic spaces afforded by punctuation points. Their conclusion certainly suggests that readers are alert for spaces, just as readers of alphabetic languages are guided by word spacing. Just as English readers, for example, are taking a 'parafoveal preview,' so too, perhaps, are Chinese

readers, but, since Chinese words may contain more than one character, a space between characters may be a link rather than a break, and thus guide the parafoveal preview. On the other hand, a space at a punctuation point in Chinese will provide a syntactic break, which provides entirely different information.

### **2.5.5 Progressive and regressive saccades and skipping in Chinese reading**

Yang and McConkie (1999) did find some similarities between their Chinese reading subjects and subjects of several previous reading studies in alphabetic scripts. They found that regressive saccades were shorter on average than progressive saccades, and that there was refixation and low frequency words were refixated more often than high frequency words. Skipping, see Section 1.2.3, tends to occur with high frequency words more often than with low frequency words, and longer words (words of two or more characters) tend to receive longer gaze duration (Yang and McConkie 1999: 218).

The effect of character complexity and the relationship between character complexity and frequency has two effects, according to Yang and McConkie: more complex words take longer to identify and can be identified less far into the periphery of vision (Yang and McConkie 1999: 218). This would account for the wide range of saccade lengths encountered in the study.

Apart from the difference in what is provided in peripheral vision, Yang and McConkie came to no firm conclusion as to why Chinese readers' saccades and fixation vary so dramatically and seem to be neither word nor character based. My own view, based on my data as described in Chapter Five, is that readers may be controlled by linguistic information, rather than visuo-spatial information, and that they are looking firstly at the composition, or headedness, of multi-character words, and having identified the word, they



are then looking at the syntactic context of the word, both of which strategies might require fixation on spaces between characters and would certainly require saccades of varying lengths.

The results of the studies of both Sun and Feng, and Yang and McConkie, appear to contradict both the view that Chinese is read with pursuit movements, and the view that it is read with a smooth glide of regular small saccades.

### **2.5.6 Saccades, fixation and linguistic information**

Inhoff, Liu and Tang (1999) take as their starting point in eye movement studies the close relationship between letters and pronunciation in an alphabetic script, and what they regard as the 'poor' relationship between the graphemic structure of a Chinese character and its pronunciation. They note that readers of alphabetic scripts 'complete the lexical analysis of a word - including the assignment of word meaning - while it is fixated' (Inhoff, Liu and Tang 1999: 232). Ease of linguistic analysis appears to determine gaze duration, as long, high frequency words receive shorter gaze duration than equally long low frequency words - a finding which corroborates those of Yang and McConkie, above.

Inhoff et al. found in their study of Chinese reading subjects a modal fixation duration of 220 ms, but also more common brief duration fixations of 160 ms or less. They also found that regressive saccades were fewer than in readers of alphabetic script (4% of saccades in Chinese readers, 15% of saccades in alphabetic readers). They claim that, like readers of alphabetic scripts, readers of Chinese determine the meaning of a character while it is fixated.

Inhoff et al. note that in alphabetic reading two types of information are provided by parafoveal preview: visuo-spatial cues and linguistic

specifications, and that linguistic information is obtained after a word is fixated. They state that the acquisition of useful information from parafoveally visible text is also required for reading Chinese (Inhoff, Liu and Tang 1999: 232). They found that in their Chinese reading subjects the perceptual span encompassed the fixated character, one to two characters to its left, and up to three characters to its right, a span which gives the reader useful information from words in the immediate context of the fixated word or character. This may be of relevance to both short term or working memory (Zhang and Simon 1985) and idea units (Chafe 1980), as phrases and clauses in Chinese tend to be short - a clause between punctuation marks or low density high frequency characters is often as short as 6 characters and therefore coincides with the perceptual span.

Their findings show that character and word meaning are more important for saccade planning than visuo-spatial layout. They argue that phonological information is of little or no use for character recognition. They conclude, crucially, that ease of linguistic analysis determines individual fixation duration, and analysis begins before fixation, (ie when the target word or character comes into peripheral vision) is carried out during fixation, and may be continued after fixation, thus extending across several fixations (Inhoff, Liu and Tang 1999: 233). Thus, they demonstrate that at perceptual level, Chinese reading does not differ markedly in nature from alphabetic reading, as has previously been supposed.

In English reading, fixation is normally to the left of centre of a word. This may reflect headedness in English - the reader may be fixating a boundary between stress (the head of the word or phrase) and unstress - the morphological information. In Chinese reading this type of information might be similarly sought; the reader may look for headedness, likely to be located in lower frequency, higher density, more complex characters, and may look

for grammatical information, likely to be located in higher frequency, lower density, less complex characters, or in zero manifestations, such as arguments and conjunctions.

Eye movement studies of both English and Chinese reading, as exemplified above, seem to support the view of Chafe (1980), in providing evidence of focus of consciousness in reading, as discussed in Section 1.5. High density, low frequency words or characters are fixated, and low density, high frequency words or characters provide syntactic, predictive information within the parafoveal view. As suggested above, fixation and saccade together may encompass both the 'idea unit' as suggested by Chafe (1980) and short term memory span as suggested by Zhang and Simon (1985).

### **2.5.7 Characteristics of Chinese text which may affect reading at syntactic level**

There are some unique characteristics of Chinese text which are well known and over the centuries have provoked debate resulting in a range of assumptions about its relative difficulty or ease and its relative transparency or opacity. Some of these assumptions verge on the mythological and need to be viewed in a more objective fashion. For example, Chinese is often said to be a 'reader responsible' language (Connor 1996), tending to express ideas more implicitly, which demands more cognitive effort from the reader as opposed to a 'writer responsible' language (such as French<sup>6</sup>) which tends to express ideas more explicitly and demands less cognitive effort from its readers. The implicitness of Chinese text is largely due to:

1. empty words;
2. null arguments;
3. the lack of inflection.

---

<sup>6</sup> French pronouns may be just as opaque as Chinese pronouns, for example 'I...' and 'lui..'

These characteristics are as important to reading as recognition of the forms of the characters.

### 2.5.7.1 Empty words and full words

It should be noted here that traditional Chinese grammar has two categories of words: ‘empty’ words and ‘full’ , or ‘solid’ words. Full words are what may be counted as nouns, verbs, adjectives and (some) adverbs. They can be used independently and have semantic content. Empty words are the function words - connectors and particles such as DE 的 (possessive, modifying marker) LE 了 (perfective particle) BU 不 (negative) and many more. There is of course, overlap, particularly of adverbs, depending on the type of text and discourse and on the view of the researcher. Just as in other languages, readers may ‘skip’ these apparently unimportant items, even though they constitute the mortar of the whole text. Wong (1995), writing on the difficulties of translation points out that ‘as there are no morphological changes in Chinese .... word order and the use of ‘empty words’ are critical, because they are the basis of word formation and sentence construction’ (Wong 1995: 209). Noting that empty words are a ‘tremendous challenge’ to the translator, Wong gives examples not only of empty words, but also of cases where empty words have been legitimately and grammatically omitted, as follows:

12.

他是 教师 我 是 学生  
 ta shi jiaoshi wo shi xue sheng  
 he is teacher I am student  
 He is a teacher but I am a student.

It is possible, but not necessary, to join these two clauses with an additive connective, a contrastive connective, a full stop, a comma or nothing. In whatever way the clauses are connected, the interpretation will depend on

the context. (There is some similarity here with English, as ‘and’ and even ‘but’ may be ambiguous or seen to be in free variation with a full stop, thus having a somewhat inaccessible meaning (Blakemore 2000).

Thus the all important, but negligible, mortar of the Chinese text may not even be present for the reader to skip. For example, connectors such as ‘and’ and ‘but’ may be omitted in an environment of parallel clauses, as above, and in a conditional sentence in Chinese RUGUO 如果 (if) is possible but not obligatory. The connector ER 而 may be additive (‘and’) or contrastive (‘but’).<sup>7</sup>

Traditional Chinese approaches to reading and cognition may encourage this neglect of empty words. There is a widespread opinion that Chinese characters link directly to thought via the right hemisphere in contrast to graphophonemic orthographies which take a circuitous route to thought via the left hemisphere (Shen 1988). (Some western researchers, Goodman and Smith among them, have argued that a written word acts as a visual image which provides direct access to the lexicon.) If this is the case, empty words, having no visual semantic content, are likely to get left behind. When a reader does skip empty words, they are tantamount to gaps in the text, just like null arguments or zero pronouns, and there is evidence of this in the translations of inexperienced readers. Pollard (1995), writing on modal adverbs ( a category of empty words) states that ‘empty or functional words have formed the backbone of the grammar for centuries’ but at the same time points out that a native speaker ‘looks upon the word as truly ‘empty ‘ and therefore not worth bothering with’ (Pollard 1995: 218). This could be viewed either as

---

<sup>7</sup> This optionality in the use of ‘empty words’, coupled with a lack of inflection in Chinese has triggered some intense debate. Some Western scholars, for example Bloom (1981), have endeavoured to show that, because Chinese has no ‘subjunctive’ or other inflectional device to express counterfactuals, and avoids the equivalents of ‘if’ and ‘whether’, Chinese speakers have less resources to think in counterfactual terms. Other researchers eg Hoosain (1991) and Au (1992) have countered this. Bloom’s position may be due to ignorance or incomplete knowledge of the language.

careless decoding at syntactic level, or as a 'top down' approach. Traditional memorisation methods used in education in Chinese speaking areas may aggravate this tendency. As individual characters and words are highly context dependent it is necessary to remember word groups and text. At the same time, the economy of Chinese text provides key words without the frills - it may be just as easy to memorise Chinese text in its entirety as it is to sift out extraneous words and focus on key words in an English text to use as memory triggers.

Tan (1995) notes the difference between an English sentence and a Chinese sentence. An English sentence is 'long, complicated and fraught with modifying elements... (its) construction is mainly through formal markers like prepositions and connectives..' A Chinese sentence, by contrast, is 'a string of short sentences, one following another according to the proper order and a logical semantic sequence...(its) construction is mainly through word order, logical sequencing of elements, resulting in syntactic conciseness' (Tan 1995: 482). Shi (1989) in his discussion of the topic chain (reviewed in Section 4.2.2.2) notes that a topic chain, or sentence (S) may simply be a series of co-ordinated clauses with zero co-ordinating conjunctions. A Chinese sentence is rich in content words, poor in function (or empty) words, or may even have many unrealised function words. Tan's notion of short sentences strung together and Shi's notion of a series of co-ordinated clauses echo Chafe's (1980) idea units, and may relate to saccades and short term memory.

This emphasis on the recognition of the full word and the backgrounding of the grammatical links is fairly typical of approaches to Chinese, among both native Chinese and Western scholars. It is hardly surprising, then, that young, inexperienced readers of Chinese may put all their effort into reading the characters and particularly the 'full words' without giving sufficient thought to the gaps and the 'empty words.' This notion is supported to a

certain extent by the eye movement studies cited above, which found that readers were more likely to skip low complexity high frequency characters and more likely to fixate high complexity, low frequency characters (see Section 2.5.5).

### 2.5.7.2 Pronouns and null arguments

There is some evidence in my data that inexperienced readers of Chinese treat pronouns as if they were empty words, that is they may ignore them, so that even a pronoun may, in effect, become a null argument in the perception of the reader. This is further discussed in Chapter Five.

Chinese is a 'pro-drop' language: pronouns are used in anaphoric reference but usually for reasons of highlighting, change of topic etc. The normal form of anaphoric reference is zero. Null arguments or zero pronouns in Chinese are discussed in detail in Chapter Four.

### 2.5.7.3 Lack of inflection in Chinese

As noted above in Section 2.1, Chinese does not have the rich inflectional systems of many European languages, and moreover tends to ellipsis and pro-drop, and tends to omit connectors. A more detailed account of the Chinese language is given in Section 4.2.

### 2.5.7.4 A note on punctuation

There is no equivalent of a capital letter in Chinese writing, as follows:

13.

|  |               |        |          |     |         |         |
|--|---------------|--------|----------|-----|---------|---------|
| 我  | 哥哥            | 叫      | 建华       | 在   | 北京      | 工作      |
| wo   | gege          | jiao   | Jianhua, | zai | Beijing | gongzuo |
| my   | elder brother | called | Jianhua  | in  | Beijing | work    |
| My elder brother is called Jianhua. He works in Beijing. |               |        |          |     |         |         |



In texts intended for use in schools a concession is sometimes made by underlining proper names - an aid of inestimable value to inexperienced readers, but unfortunately not widely available.

It is important to note the role of punctuation in Chinese syntax. Traditionally, prior to the introduction of texts in vernacular Chinese ('bai hua wen') punctuation was not used in Chinese texts (Hoosain 1991: 22), and was only introduced as a result of Western influence. Taylor and Taylor claim that the use of punctuation is not as consistent and regular as in the West (Taylor and Taylor 1995: 107). But modern readers of Chinese do have the luxury of commas, semi-colons, colons, and full stops.

Full stops are often used where a European reader would expect a colon or comma, for example between a 'because' clause and its 'therefore' clause. Commas are often used where a European reader would expect a full stop. Sometimes the sentence initial subject or topic is marked by a comma.<sup>8</sup> There are certain pragmatic conventions, but there does seem to be a degree of free variation. A bilingual or non-native reader of Chinese may therefore experience some difficulty in parsing between and around punctuation marks. As noted previously, the only true physical spaces in Chinese text coincide not with word boundaries, but with punctuation marks or clause or phrase boundaries, giving rise to the regularity highlighted by Chen in his quantization process, described in Section 2.4.1.

Inhoff and Liu (1997) point out that punctuation marks in a Chinese text, 'though quite small, often occupy a full character space' and probably play a vital role in signalling clause boundaries in a writing system which has few syntactic cues otherwise (Inhoff and Liu 1997: 263).

---

<sup>8</sup> The Chinese word for sentence is *ju* 句 and appears as *juzi* 句子 (sentence), *zhuju* 主句 (main clause) and *congju* 从句 (subordinate clause) etc.

## **2.6 Potential problems of comprehension of Chinese text**

### **2.6.1 Potential problems of low level decoding of Chinese text**

Given the paucity of grammatical information in Chinese text, as compared with, for example, highly inflected French or Russian, the competent reader needs to concentrate efforts on firstly, accessing the characters accurately, and secondly, exploiting the text for inference. Traditionally, Chinese characters are learned by rote and practice. As explained above, it is possible to 'work out' the pronunciation and meaning of some characters but generally, they must be memorised before any meaningful reading is possible. If a character is mistaken for one that sounds the same, or looks very similar, but has a different meaning, the reader may adjust his or her schema to suit the mistake and thus achieve a coherent, but erroneous comprehension of the text.

### **2.6.2 Grammatical /syntactic sensitivity in reading Chinese text**

However accurate a reader's comprehension of vocabulary is, the syntactic code which constrains the order and use of words in sentences, defining the functional and formal properties of those words, must be understood. A linguistically aware reader who is able to make these distinctions is better able to store and retrieve information when reading.

Grammar may be much more of a problem for a reader than vocabulary. For example, Cohen, Glasman, Rosenbaum-Cohen, Ferrar and Fine (1988) found that Hebrew speaking readers of English had difficulty with complex 'heavy' noun phrases. They were helped by basing their parsing on the head of the NP or VP. This kind of grammatical coding is automatised by native or expert readers and only carried out step by step by inexperienced readers, or when overly complex (Cohen et al. 1988). Inflected languages ease grammatical decoding considerably by the agreement systems which link nouns, adjectives and verbs. This is borne out by results of fixation studies, as

discussed in Section 2.5. As explained previously, Chinese lacks these morphological clues.

Grammatical sensitivity has been found to correlate positively with reading ability as reported by Bryant (1995). However, none of the tests reported by Bryant were able to distinguish between semantic and grammatical aspects of language. The kind of grammatical sensitivity test used with Western languages may not be possible with Chinese because of its lack of inflection.<sup>32</sup>

### **2.6.3. Empirical studies of grammatical /syntactic sensitivity in reading Chinese text**

#### **2.6.3.1. Timing and comprehension**

Theories of the timing of comprehension in reading are represented at two extremes of the spectrum by the following views:

1. That comprehension of a word is delayed while readers wait for context;
2. That comprehension is instantaneous.

H.-C. Chen (1999) notes that, with regard to alphabetic reading, intermediate views are now more generally accepted. He carried out violation experiments to determine whether this would also be the case with reading Chinese. (However, as noted above in Section 1.4.8 on inference, readers of English probably tend to delay drawing inferences until it is necessary, or until they have sufficient information.)

In this type of experiment, critical words are replaced by lexical, semantic and syntactic-semantic violations. Patterns of disruption in readers' reactions are analysed. An example of Chen's violated texts:

14.

每月总有两三次，全校学生列队在图书馆集合，总有一个或两个坏孩子，会给推进隔壁房间，鞭打（学裕，鞭打著，量度）得皮破血流，其余的人坐在那里，哆嗦著听他们呼号尖叫。我恨透了这所学校，那两年多提心吊胆的日子，真不知道是怎麼过的。

Two or three times a month the whole school was marshaled (sic) in the library, and one or more delinquents were haled (sic) off to adjoining apartment and there flogged (seprool, flogging, measured) until they bled freely, while the rest sat quaking, listening to their screams. How I hated this school and what a life of anxiety I lived there for more than two years.

(Chen 1999: 194)

Chen found that all three types of violation caused interference at the end of the sentence, showing that readers tried to use context to resolve the violation. In the immediate vicinity of the replaced word, however, lexical violations produced an immediate disruption in reading, whereas semantic-syntactic violations caused disruption at the word+1 position. With semantic violation, disruption occurred three words after the critical one. Chen found that a considerable amount of contextual information was needed to detect the semantic violation of a verb. His explanation is that

*'since the main function of a verb is to relate the nouns in a sentence in order for a reader to comprehend that the verb provided in the semantic violation condition is anomalous, a reader has to know what the nouns of the sentence are' (Chen 1999: 265).*

This would suggest that the reader of Chinese text uses context at more than one level in reading, and is comprehending immediately and after confirmation through context. This supports the findings of Inhoff et al., discussed above, in that it is complementary to the notion that linguistic

information rather than purely visuo-spatial information guides the reader. The notion can be extended to cover the sentence, as well as the perceptual span (Chen 1999: 265).

### 2.6.3.1.1 The role of the verb in comprehension

The delay in disruption only occurred when the semantically violated verb was in an SVO sentence. When, however, the semantic violation of the verb occurred in either SOV or OSV sentences, using BA 把 and BEI 被 to modify the word order, the violation was detected quickly at word+1 position (H.C. Chen 1999: 266). Examples of SVO, SOV and SAV/OSV sentences are given below:

15. SVO        他打我  
                  ta da wo  
                  he hit me  
                  He hit me.
- SOV        他把 我打  
                  ta ba wo da  
                  he BA me hit  
                  He hit me.
- OSV/SAV    我 被他打  
                  wo bei ta da  
                  I (passive) him hit  
                  I am/was hit by him.

It is important to note here that OSV type sentences could be construed as SAV (subject-agent-verb) sentences. Moreover, conversion to SOV and OSV/SAV type sentences it is not simply a question of word order. The use of BA 把 not only highlights the object, but also introduces an element of

definiteness. The use of BEI 被 not only conveys passivity, but also of unpleasantness.<sup>9</sup> In both cases the topicality of the sentence is altered. Chen himself notes that both SOV and SAV/OSV sentences are marked, whereas the SVO sentence is unmarked.

Chen concludes that the semantic processing of a verb relies on the positioning of the main participants (nouns and verbs) in a sentence and is affected by word order. He notes that word order is important for certain types of analysis, such as integration, and that verbs 'play a more important role than nouns in sentence integration in Chinese' (Chen 1999: 267). What Chen does not tackle is the question of comprehending the verb in an environment of null arguments or zero pronouns. This relative importance of verbs in anaphoric resolution of zero pronouns will be discussed again in Chapter Four.

#### **2.6.3.1.2 The role of the noun and classifier in comprehension**

Chen repeated these violation experiments with sentences in which the critical words were nouns (eg 'tea') or classifiers (eg 'a cup (of tea)') or nouns modified by classifiers, and found similar effects. Noticing that dramatic disruption was caused by violations which featured 'empty' words (*ba*, *bei*) playing an important grammatical role, Chen related these findings to the traditional categories of full and empty words and attempted to determine whether these traditional categories could have any psychological implications (Chen 1999: 269).

The results of the experiments taken together suggested that when nouns and classifiers are violated either syntactically or semantically, the patterns of disruption showed that both these types of content words contribute equally to both syntactic and semantic operations in reading Chinese. Chen does not

---

<sup>9</sup> This now seems to be changing - Chinese passives are increasingly used in a neutral sense.

refer to the elliptical and anaphoric role played by classifiers, but this study may suggest that it is easier to read and understand 'full' or content words than 'empty' or function words.

With verbs, the syntactic violations produced stable disruption, but the semantic violations produced differing disruptions depending on the position of the verb in the sentence. When BA and BEI occurred in the disrupted sentences, syntactic violations produced reliable and long lasting effects, suggesting that the linguistic properties of various types of words have psychological implications (Chen 1999: 273), a finding that supports the view of Inhoff et al. that eye movement behaviour is determined by linguistic information rather than visuo-spatial cues. According to Chen's study, the patterns of delay in reading disruption caused by violation (immediate for lexical violation, word+1 position for syntactic violation and word+1 or more position for semantic violation) do not support the extreme delayed processing hypothesis, or the radical immediate processing hypothesis (Just and Carpenter 1980). This is probably because linguistic information is not overtly provided in the form of inflectional morphology but must be inferred from the text itself through, for example, word order. Chen concludes that Chinese reading subjects rely on a context-dependent, diffused reading strategy rather than a focused word-dependent strategy more likely to be used by readers of alphabetic systems (Chen 1999: 273).<sup>10</sup>

Chen's study demonstrates unequivocally, that in reading Chinese text lexical access precedes any other form of comprehension analysis. This result contrasts with the view (exemplified by Rayner and Pollatsek 1989) that in English reading syntactic and semantic parsing is simultaneous, or 'driven by

---

<sup>10</sup> Chen sometimes uses the term 'alphabetic language,' sometimes 'alphabetic system,' apparently not distinguishing alphabetic script from inflected languages. This is a crucial distinction which ought to be made. Japanese and Korean, for example, have inflected verb endings but do not use alphabetic systems. They do, however, use phonologically based syllabaries. Bahasa Indonesia, for example, uses an alphabetic system but has little inflection. A focused, word-dependent strategy probably depends more on the morphological structure of a language than on its orthographic system.



higher order information' (Rayner and Pollatsek 1989). The issues of syntactic and semantic analysis in reading Chinese text are less clear. Chen concludes that the verb plays an important role in integration in reading comprehension, an idea supported by research in English reading (Fillmore 1971, Schank 1975). However, his results show that the Chinese reader relies more heavily on the noun in comprehension analysis. Obviously, this notion falls short when null arguments or zero pronouns, so common in Chinese, are considered. The importance of the verb in comprehension of Chinese text will be addressed more fully in Chapter 4 when the null argument is discussed.

Chen suggests that classifiers exist to support the processing of nouns. Presumably he has in mind the elliptical role of the classifier, but does not expand on this. If this is what he intends, it is still consistent with the traditional idea of full and empty words. A classifier, or number+classifier standing in place of number+classifier+noun constitutes a kind of attenuated or reduced anaphor, similar to the role of a pronoun, and plays a role in anaphora. But it is not quite zero, as a classifier will always be preceded by some kind of determiner such as a number or demonstrative.

In Chinese, as in other languages, the noun phrase may carry three degrees of information - full information in a noun, partial information in an anaphoric form such as pronoun or classifier, and zero information in a null argument or zero pronoun. In some languages verbs may also show degrees of information and take on an anaphoric role, as for instance in English when auxiliaries stand in place of full verb forms, for example:

15. 'Did you lock the door? Yes, I did.'
- |            |      |          |            |
|------------|------|----------|------------|
| 锁了         | 门    | 没有?      | 锁了。        |
| suo le     | men  | mei you? | suo le.    |
| lock (asp) | door | NEG ?    | lock (asp) |

Did you lock the door? Yes, I did..

In Chinese, however, full verb forms are usually required. The information load is borne mainly by verbs or predicates, rather than by nouns or arguments. This applies in spoken Chinese and in written texts. It is possible, as Chen suggests, that Chinese reading subjects rely initially on the noun for comprehension, but require information from the verb to achieve integration in comprehension. There are cases of verbless sentences in Chinese and this phenomenon will be dealt with in Section 4.2.1.2.

### **2.6.3.2 Cue validity or information value of linguistic forms in reading**

#### **2.6.3.2.1 Animacy and word order as cues in comprehension**

The roles of animacy and word order in comprehension of written and spoken Chinese were investigated by Miao (1981, 1999). While Italian speakers, using a highly inflected language, determine the subject of a sentence by means of lexical meaning and morphological markers, English speakers, using a language with some inflection, but far less than Italian, determine the subject mainly by word order (Miao 1999: 280). Asking subjects to act out presented sentences using toys, Miao (1981) found that while word order played some part in determining the subject or agent of the sentence animacy had a greater effect. (Miao does not distinguish clearly between subject and agent).

The sentences presented to subjects in Miao's study all consisted of a verb and two nouns, the two nouns alternating between animate and inanimate, and the structure of the sentence varying between NVN, VNN and NNV, as in Chen's study described above. Miao found that when the information of word order conflicted with semantic information, word meaning or lexical constraints predominated in determining the subject or agent of the sentence (Miao 1981). Reviewing this study in 1999, Miao points out that the findings

are contrary to expectations, as many people believe word order to be paramount in Chinese syntax, but in fact, word order in Chinese 'is far more variable, and much less powerful than was assumed' (Miao 1999: 283). Miao concludes that event probability is the essential cue for sentence comprehension in Chinese.

Li, Bates, Liu and McWhinney (1992) cite the studies carried out by Miao which demonstrate the importance of animacy for a Chinese speaker in determining the agent of the sentence and also Miao's work which found evidence of an interaction between word order and animacy. They also cite supporting work by Chen, Tzeng and Bates which found that animacy had a stronger effect than word order and similarly found an interaction between word order and animacy (Chen, Tzeng and Bates 1990).

Li et al. based their similar experiments on the competition model - 'an interactive model of language use in which the cue validity or the information value of linguistic forms in a given language plays a probabilistic role in the process of mapping between surface forms and underlying functions' (Li et al. 1992: 208). Li et al. note that the strength of the relationship between form - grammatical devices and semantic cues - and function - categories such as agent and topic - varies from language to language (Li et al. 1992: 208).

Li et al. define cue validity as the product of how often the cue is available, and how often it assists the reader to the right answer. For example, word order in English has greater validity as a grammatical cue than morphology, whereas morphology in Italian has a greater validity than word order.

They looked at the effect of animacy as a cue in Chinese within four possible configurations - SVO, OSV (or SAV), SOV and VOS - noting that only the object can occur post-verbally, and that the pre-verbal position is not particularly associated with a definite function. This is a characteristic which is very different from English, for example, which is strongly inclined to pre-

verbal positioning of the subject , and although post-verbal positioning may indicate object in English, this is not such a strong cue as pre-verbal positioning. Li et al. note that not only the subject, but also other constituents of the sentence may be omitted in Chinese 'as long as the context provides clues as to who does what to whom' (Li et al. 1992: 212). This, of course, is the crux of the problem. Whatever is said about topics, agents, subjects, or objects etc. in Chinese must be qualified to make allowance for the effects of the ubiquitous null argument or zero pronoun. Li et al. indeed admit that 'omission reduces the reliability of word order cues' (Li et al. 1992; 212) and it must also reduce the reliability of animacy cues.

*'Sentence roles are not determined solely by grammatical devices. It (sic) can also be influenced by the semantic properties of the noun phrase itself, such as animacy, ie whether a noun phrase indicates an animate (including humans and animals) or an inanimate object. Comrie (1981) discusses in detail the interrelation of animacy with other syntactic and semantic factors, eg number, gender, case and marking, showing that animacy is relevant and important to grammatical distinctions' (Li et al. 1992: 213).*

Li et al. go on to suggest that certain verbs, in Chinese as well as in other languages, have a semantic bias which will act as a cue in comprehension (Li et al. 1992: 213). Perhaps particularly so when topic, agent, subject and object may be absent from the surface form of the sentence.

On the basis of the theory of cue validity, Li et al. asked Chinese speaking subjects to match pictures to sentences in the following conditions:

animate - verb - animate (AVA)

inanimate - animate - verb (IAV)

verb - animate - inanimate (VAI)

(There was no inanimate - inanimate condition.)

The subjects were asked to determine who the ‘actor’ was, and were found to rely primarily on animacy, but Li et al. found that word order interacted with animacy. They conclude that because of the ‘impoverished’ morphology of Chinese, Chinese speakers need to make use of all possible cues and integrate them’ (Li et al. 1992: 229). This gives support to H.-C. Chen’s view that Chinese readers employ a more diffuse reading strategy, see Section 2.6.3.1.2.

#### **2.6.3.2.2 Syntactic and semantic information as cues in understanding pronouns**

Miao and Song (1995), in an investigation of the comprehension of potentially ambiguous pronouns, asked subjects to complete sentence fragments ending in a pronoun, for example (when both proper names refer to males):

16.

王 表 批评 张信，他。。。。

Wang Biao pi ping Zhang Xin, ta.....

Wang Biao criticise Zhang Xin, he....

Wang Biao criticised Zhang Xin, he.....

(Miao and Song 1995)

They found that assignment of the pronoun to either the first or second noun in the initial clause depended largely on the verb - certain verbs had an implicit meaning of causality (Miao 1999: 286), a view supported by Chen and Hung (1995). Where the first (complete) clause had no explicit link, such as ‘because,’ subjects seemed to use a parallel function strategy, that is, the pronoun refers to a noun in the preceding clause that has the same grammatical function, for instance subject. Miao and Song found that where there was no explicit conjunction to link clauses (as discussed in Section 2.1.1) subjects were most likely to use the parallel function strategy regardless of the meaning of the verbs concerned (Miao and Song 1995).

In further studies of pronoun comprehension, Miao (1996a, 1996b) presented Chinese reading subjects with three types of sentences. In the first type, the pronoun in the second clause referred back to the first noun, or subject of the first clause, and the information for disambiguation of the pronoun was carried entirely in the verb of the second clause:

17.

小 梁 终 於 找 到 了 迷 路 的 小 孩 , 他 谢 绝  
 Xiao Liang zhongyu zhaodaole milu de xiao hai, ta xiejue  
 Xiao Liang at last find (asp.) lose way (mod.) small child, he refuse  
 接 受 一 笔 报 酬  
 jieshou yibi baochou.  
 accept one (measure) reward  
 Xiao Liang finally found the lost child, he refused to accept a reward.

In the second type, the pronoun referred back to the second noun in the first clause, the object, and again disambiguation was carried entirely in the verb of the second clause:

18.

经 历 在 办 公 室 当 中 表 扬 了 小 钱 , 他 挽 回 了  
 Jingli zai bangongshi dang zhong biaoyangle Xiao Qian, ta wanhuile  
 manager in office in centre praise (asp.) Xiao Qian, he retrieve  
 (asp.)  
 一 大 笔 损 失  
 yi da bi sunshi  
 one big (measure) loss  
 The manager , in the office , in the presence of all, praised Xiao Qian,  
 he retrieved a great loss.

In the third type of sentence the pronoun was ambiguous because the verb provided no disambiguating information:

19.

小 梁        终 於    找 到 了 迷 路 的    小 孩 ,        他 当 时    累 得  
 Xiao Liang zhongyu zhaodaole milude    xiao hai,        ta dangshi leide  
 Xiao Liang at last    find (asp.) lost way small child, he that time tired  
 站 立 不 住  
 zhanlibuzhu.  
 stand not firm  
 Xiao Liang finally found the lost child, he was so tired he could not  
 stand. (Miao 1999: 287)

Miao found that subjects invariably looked to the verb to aid in processing the pronoun. Subjects used the parallel function strategy with the third type of sentence where there was no obvious semantic bias (Miao 1996a). These experiments were based on single sentence stimuli. It is possible that where a 'garden path' or ambiguous sentence occurs in a text, readers will look at other parts of the text, or their own background knowledge, for disambiguation.

Miao also found that proximity of the pronoun to its antecedent only affected processing in sentences of three clauses, not in sentences of two clauses. Thus it seems that semantic information is more important in the Chinese reading process than syntactic information. Miao's investigations on pronoun processing are germane to the discussion of zero anaphors and null arguments in Chapter Four.

Chen and Hung's (1995) similar experiment, however, showed that Chinese reading subjects benefited considerably from structural factors to guide their comprehension, rather than semantic information. Miao attributes this discrepancy to the different experimental methods used (Chen and Hung



used the 'moving window' technique) and modestly indicates that on-line experiments may be preferable (Miao 1999: 290). But it is possible that on-line experiments cannot fully and authentically represent the reading process, only a certain very limited kind of reading process.

Peng and Liu (1993) used garden path type sentences, as follows:

20.

王太太 邀请 了林小姐 担任 晚会 的

Wang Taitai yaoqingle Lin Xiaojie danren wanhui de

Mrs. Wang invite (asp.) Miss Lin responsible evening party (mod.)

表演 嘉宾

biaoyan jiabin.

performance honoured guest

(1)她 很 兴奋 因为

(1) Ta hen xingfen yinwei

(1) she very excited because

(2)林小姐 自认为 是 城里 最有名的 歌星

(2) LinXiaojie zirenwei shi chengli zui youming de gexing.

(2) Miss Lin herself think is city most famous (mod.) singer

Mrs. Wang invited Miss Lin to be the honoured performing guest at the evening party, (1) she was very excited because (2) Miss Lin considered herself the most famous singing star in the city.

(Peng and Liu 1993: 291)

They found that semantic information was paramount in sentence comprehension and that semantic constraints affected syntactic analysis (Peng and Liu 1993).

### **2.6.3.3 Linguistic information affecting reading for comprehension and reading for recall**

H.-C. Chen (1992) used violation experiments to test the strength of animacy in relation to the absence, or minimal presence of syntactic information in Chinese. Using whole texts, with violations embedded therein, he found that Chinese reading subjects relied most heavily on animacy, but also on other cues such as word order. He reiterates the assertion that functional characters indicating syntactic information (ie empty words) are not always present in sentences. He suggests that syntactic factors may not be crucial for comprehension (Chen 1992: 180).

Chen's method in this study was to use the 'moving window method' as later used by Sun and Feng (1999) and described above. The text was presented on a computer screen, one word at a time, the context being represented by strings of dashes to represent words which have already been read and which have not yet been read. The subject is in control, pressing the key for a new word when ready to continue. The reading time is recorded and measured (Chen 1992: 182).

Chen claims that this method 'closely resembles a conventional reading method' and provides data comparable to those obtained in eye movement monitoring (Chen 1992: 182). To be fair, the programme was modified to allow for retrogression in the text, but it does not present whole text as in a normal reading situation. However, Chen used whole texts, rather than single sentences. The texts were Chinese translations of English texts taken from Time and Newsweek. After reading the whole text on line the subjects, native Chinese speaker/readers, were divided into two groups, one asked to recall the text, and the other to answer forced-choice comprehension questions.

Chen found distinct differences in reading times of characters between the two groups. Those reading for recall showed longer reading times for

characters at the ends of clauses, when new concepts were introduced, when new concepts in a sentence increased in number, and at the beginnings of lines. Reading times were shorter for this group when characters were less complex and of higher frequency (Chen 1992: 188). Chen claims that these results show a syntactic or structure-oriented strategy.

Those reading for comprehension were strongly affected in their reading by three variables - the number of new argument nouns in the sentence, a new concept, and the end of the line. Chen concludes from this that when reading for comprehension Chinese readers use a semantic or reading oriented strategy. (Chen does not define his term 'new concepts' and does not explain the relationship between 'new concept' and 'new argument.' An abstract argument could be a concept. This leaves a gap in Chen's results, as it may well be the case that a new concept cannot exist without a new argument, or vice versa, in which case there could be more similarity between the two groups than is at first apparent.)

In this experiment Chen found that neither visuo-spatial properties such as complexity, nor frequency, had any effect on reading times in either subjects reading for comprehension, or in subjects reading for recall. He concludes that Chinese reading subjects 'pay less attention to individual sentence units....and more attention to the ends of syntactic or physical boundaries' (Chen 1992: 192). This would suggest that Chinese readers are looking to gaps, or boundaries for guidance, in the same way as English readers look to gaps between words.

#### **2.6.3.3.1 Linguistic information affecting Chinese and English reading**

Certain of one significant variable in this experiment, that is, the new argument nouns, but uncertain of whether their effect was causal, Chen (1992) carried out violation task experiments using 40 narrative passages from

the Chinese edition of *Readers' Digest*, on average 83 characters long. The same passages, in English, were provided for English reading subjects.

The comprehension scores for the Chinese reading group and the English reading group were comparable, but the patterns of disruption caused by violation were distinct. While English reading subjects experienced dramatic disruption at the critical word with all three types of violation - syntactic, semantic and lexical, the Chinese reading subjects experienced disruption at the critical word only when the violation was lexical. With semantic and syntactic violations, the disruption was later, at the end of the clause. Chen concluded from this that 'English readers .... use a more focused, word dependent strategy than do Chinese readers, probably because various aspects of information carried by individual words in English are rather informative' (Chen 1992: 198).

The English reading subjects in the experiment were Chinese background bilingual subjects.<sup>11</sup> Chen takes this as evidence of support for the theory of the concept-mediation hypothesis, that is, the two 'languages of proficient bilinguals are operated independently, and information in the two languages is not directly or strongly associated, but connected through an amodal conceptual system' (Chen 1992: 199).

Further to this discussion of the concept-mediation hypothesis, Chen notes that some research supports a theory of different processing mechanisms according to different orthographic systems (Chen 1992: 200). However, this notion is largely based on the results of experiments designed to test lexical tasks, focusing on words or characters in isolation, which leads Chen to surmise that variation in orthographic systems may show differences in processing at encoding and perceptual level for simple word recognition

---

<sup>11</sup> One of the problems with Chinese reading research is that in most cases English background monolinguals are compared with Chinese background bilinguals. Studies of Chinese background monolinguals and English background bilinguals are rare.

tasks, as discussed in Sections 0.6 and 2.1.2. Complex reading tasks, such as sentence interpretation and text comprehension, on the other hand, are dealt with at a higher level of processing. By way of example, Chen notes that in his experiments subjects show a longer reading time for characters in technical passages than in prose passages (Chen 1992: 201).

#### **2.6.3.3.2 Thematic information in comprehension**

Siu (1988) pursuing van Dijk and Kintsch's (1983) discursual approach to comprehension, looks at the notion that searching for an important idea from a passage is an integral part of the reading process. Siu maintains that the reader converts information in a text into propositions and integrates these propositions into larger units - macropropositions - built on the reader's mental operations, which are in turn based on the semantic content of the passage. Siu investigated the way textual clues affect readers' processing of the main idea in Chinese texts (Siu 1988: 163).

According to Siu, initial mention signals thematic importance, and frequent mention and concluding remarks or summary items signal thematic importance to a lesser, or secondary extent (Siu 1988: 163). This notion parallels that of topicality which is discussed in Chapter Four. Following van Dijk and Kintsch (1983) Siu maintains that readers will devote 'keener' attention to what they perceive as the theme and retain it for longer. This will vary according to the ability of the reader, and better readers are more sensitive to important statements and thematic structure, and are better able to integrate the information.

In Siu's experiment, 9th Grade school pupils in Hong Kong (around fifteen - sixteen years old) were asked to judge whether a given item ( a two character word) was the main theme of a passage. Siu found, contrary to expectation, that subjects perceived initial and terminal mention as having equal thematic

importance, a result which appeared to contradict van Dijk and Kintsch (1983). However, Siu points out that the passages in the study were shorter than previous experimental material and claims that reasoning ability was more important than processing time in the process of thematic abstraction, a rationale which would appear to undermine his position as his experimental method was to record response latencies. Additionally, Siu found no significant difference in results in readers of different abilities, but did find that his readers responded to a recurring theme marker (Siu 1988: 168).

Siu does not mention that whereas his experimental study investigated Chinese, previous studies of this nature had investigated English, in which thematic structure of a text might not be comparable to that of Chinese. Discoursal theme and syntactic topic-comment structure, (see discussion in Section 4.2.2) for example, need to be clearly distinguished. In the report of the study, the original Chinese texts are not given - only an English translation (as follows) is given, which again, undermines Siu's position, as it is not clear what, exactly, Siu is examining.

'She is known for her stubbornness (character). She never listens to others. Even if she is wrong, she will still persist. Once she has decided on something, she will firmly carry on with the task and accept no advice. We can do nothing with her character (stubbornness). (initial mentioning)

Stubbornness (main theme)

Task (distracter)'

(Siu 1988: 164)

In this example it appears that the theme is stubbornness, and the translations seem to suggest that the thematic information in the texts may have been

given in the form of stative verbs, which may act as adjectives, or even also as nouns. Thus, the discursal themes of the experimental passages may have been carried in the predicates, or the whole propositions, rather than in the arguments. If this is the case, whatever the shortcomings of Siu's study, it shows the importance of the verb in processing of Chinese text.

A further consideration of Siu's study shows that the passages consisted of the description of the attributes of a human being. The main argument, and continued topic throughout each short passage was animate, and therefore, if as suggested above, animacy is paramount in comprehension, the 'theme' (in grammatical terms the predicate, and in informational terms, the attributes) might be viewed by the reader as subordinate to the animate topic. It would then not be surprising for readers, when asked forced choice questions, simply to assign equal importance to initial and terminal mention of the theme.

#### **2.6.3.4 The application of formal schema in reading Chinese**

As described in Chapter Two, above, the proposition links sentence comprehension and text comprehension. A sentence represents one or more propositions, each consisting of argument and predicate, and the sequence of propositions in a text provides the text structure, or formal schema.

In an attempt to compare the respective applications of formal and content schema in L1 and L2 by bilingual readers, Chen Daxing (1990) used recall and multiple choice comprehension tasks, along with introspection questionnaires designed to investigate subjects' reading strategies. The subjects were Chinese native speakers carrying out advanced postgraduate studies in America and a group of monolingual English native speakers (Chen 1990).

Chen noted that previous studies had found that native speaker/readers outperformed L2 readers in recall of texts, and that native speaker/readers



remembered detail as well as gist (Connor 1984). Chen's aim was to investigate the effects of text structure, identify reading strategies and look for possible transfer of metacognitive skills across the two languages of Chinese English bilingual subjects (Chen 1990: 277).

Chen's test materials consisted of two English passages and two Chinese passages, one of each scrambled. The additional practice test was in English. The tasks were multiple choice comprehension, recall of the passages and a metacognitive questionnaire. Raters analysed the recall protocol for propositions by counting the number of propositions in each subject's recall of the texts (Chen 1990: 279).

Chen's results showed that the bilingual readers were more sensitive to text structure with English texts than with Chinese texts, and that the readers recalled more from the English texts than from the Chinese texts. (It is interesting to note that the experimenter expressly discouraged the subjects from memorising the Chinese texts - behaviour that might be intuitive in a Chinese educated academic.) Chen's results are in marked contrast to those of Connor. Not only did the Chinese bilinguals recall the English texts better than they recalled the Chinese texts, but they also recalled the English texts better than the native English speakers recalled them.

This point invites discussion. The first thing to note is that Chen points out that there may be differences in the number of propositions and arguments in Chinese and English texts (Chen 1990: 279). For this reason the recall protocol was scored for the number of sentences and the total percentage of propositions - points were not given for arguments without predicates or for predicates without arguments. As has been discussed elsewhere, the null argument is ubiquitous in Chinese text. This could mean that readers in Chen's experiment recalling a predicate with a null argument would lose a

point. This is not made clear in the report, but is a possible explanation for Chinese readers' superior recall of English texts.

A second point to note is that Chen carried out the experiments with a 'comparison' group of American native English readers. The finding that their recall was worse than that of the Chinese readers leads Chen to the conclusion that text structure has less effect when one reads one's native language than when one reads L2. It is possible that educational background gives readers very different memory ability. Moreover, conditions for the two groups of readers in Chen's experiment were very different. The Chinese bilingual readers performed the tasks individually, at times and places convenient to themselves, and were paid to do so, while the American English readers performed the tasks in a group setting, and there is no mention in the report of any payment for the American English readers: comfort, leisure, motivation and privacy have a profound effect on one's ability to concentrate.

The results of this experiment are surprising, but the explanation could lie in the experimental method of analysing comprehension of propositions. For responses in Chinese, perhaps the null argument should be taken into account, allowing predicates without arguments to be accepted as propositions. Chen's conclusion that text structure has a significant effect on recall is probably true, but needs to be modified to allow for Chinese text structure, that is, sequences of propositions and reference chains which have a base of null arguments.

The answers to the metacognitive questionnaire did not reflect the Chinese bilinguals' superiority in reading - they undervalued their English reading strategies greatly (Chen 1990: 287). The Chinese bilingual subjects also claimed in their questionnaire answers that background knowledge was the most difficult factor in reading Chinese, whereas it was not the most difficult

in reading English (Chen 1990: 287). This is not explored in Chen's report, but has profound implications for the application of content schema in reading. It may imply that background knowledge is more important in interpretation of Chinese text, which has relatively little linguistic information, than in interpretation of, for example, English text, in which there is more linguistic information, ie. inflectional morphology. This again reflects Chen's notion of more diffuse strategies in readers of Chinese. It could even account for the phenomenon recently found by Leong and Hsia (1996, reviewed in Section 4.1) in young Cantonese speakers who adopt English grammatical structures into their Cantonese speaking and writing - they may be requiring more concrete, detailed linguistic information in certain aspects of their discourse.

## **2.7 Summary and conclusion**

This thesis assumes that reading is an interactive process - as described in Chapter One - which takes account of parallel and perhaps simultaneous processing of decoding at letter and word level, lexical access, parsing skills, knowledge of text structures and application of schemata. Most reading research has investigated English, and a great deal of recent research into the reading of Chinese text has been based on this. Experimental work on Chinese has in the main been reductionist, and I have attempted to draw these studies together to create a picture of processing at all levels.

Short term memory or working memory is important in reading. Short term memory capacity for stretches of Chinese text and English text is similar, in the case of Chinese about seven non-homophonic characters, and it appears that in both languages there is an acoustical factor in memory span (Zhang and Simon 1985); Chinese readers with a higher working memory capacity name characters faster and are less dependent on inference than those with a lower working memory capacity (Cui and Chen 1997).

It seems that Chinese readers perceive groups of characters, rather than single characters one by one. While the character is perceived by many as the 'sociological word,' Chinese readers do use and recognise words, or groups of characters which act as constituents (Chen 1996). There may be a constituent superiority effect rather than a word or character superiority effect in Chinese (Mattingley and Hsiao 1997).

Syntactic and semantic processing may be revealed in the study of eye movements. Eye movements in Chinese reading consist of fixations and saccades similar to those used in the reading of alphabetic orthographies, but the saccades appear to be shorter and more regular (Chen 1996). Retrogressive saccades are used for checking and monitoring. Chinese readers tend to fixate low frequency, high density, complex characters and words - the so-called full words, and tend to skip high frequency function words - the so-called empty words (Yang and Mc Conkie 1999). However, 'skipping' may provide a parafoveal preview of linguistic information, but it may be different from the kind of information available in alphabetic orthographies, particularly where spacing and punctuation are concerned (Inhoff, Liu and Tang 1999).

Violation experiments suggest that in Chinese reading lexical access precedes any other form of analysis (Chen 1999). Chinese readers may rely heavily on the noun in comprehension analysis. Because of the lack of morphological inflection, Chinese readers may use a context dependent, diffused reading strategy, delaying decisions until more information is available.

Linguistic forms of different types may have different information value in comprehending text. For example, word order is less important in Chinese reading than traditionally thought, but other cues, such as animacy, may be of major importance in determining the agent of a sentence (Chen 1992, Miao 1999). Taken together, word order and animacy are found to interact (Chen,

Tzeng and Bates 1990). However, it appears that semantic information, particularly in the meaning of the verb, is more important in processing than syntactic information, for example in 'garden path' sentences (Peng and Liu 1993). When Chinese readers do take note of syntactic structures, they are more likely to pay attention to clause or phrase boundaries than to sentence boundaries (Chen 1992). Complexity is also an issue for the Chinese reader, for the more new arguments or concepts there are in a text, the longer it will take to read, and complexity of concepts in a text seems to have a greater effect on comprehension than visuo-spatial characteristics or effects of frequency (Chen 1992).

Apparently Chinese reading subjects seem to perceive initial and terminal mention as having equal thematic importance (Siu 1988), a finding which may have some bearing on topic persistence. In other experiments, Chinese readers appear to be more sensitive to text structure than English readers, and recall texts better than English readers (Chen 1990). Chen also suggests that background knowledge may be more important in the comprehension of Chinese text than in the comprehension of English texts (Chen 1990).

If Chinese text is read differently from languages which have alphabetic orthographies, the differences are likely to be the result of the typology of the language rather than the physical form of the orthography. Given the elliptical nature of Chinese, differences or difficulties may lie in firstly, processing of propositional structures of the text, and secondly, in the way in which schemata are applied by the reader.

The review of some of these psycholinguistic experiments into the nature of reading shows the shortcomings of reductionist experiments. For example experimenters do not distinguish between alphabetic orthographies and inflected languages, they may provide more comfortable conditions for one set of subjects than for another, they confuse grammatical subject and agent,

they neglect to exemplify their experimental texts, and so on. But the main problem with reductionist experimentation is that only tiny bits of the reading process can be studied in this way. In order to study reading it is necessary, in addition to laboratory studies, to study uncontrolled reading of whole texts in normal settings, and use subjects who are not psychology students.

It may therefore be useful either to analyse subjects' spoken or written protocols of reading, or to study translation, which is a simultaneous written protocol of reading, in conjunction with experimental methods. Translation as reading is discussed in the following chapter.

## **Chapter 3: Translation and its relationship to reading**

### **3.0 Aim and introduction**

This chapter aims to relate the theory and practice of translation, in professional and non-professional contexts, to the practice of reading, which is central to the translation process. The cibliste-sourcier dichotomy and the prescriptive-descriptive dichotomy are discussed as they are found in the Western tradition of translation. Some traditional Chinese approaches to translation - faithfulness, understanding and elegance - and some more modern Chinese approaches - directness and indirectness - are outlined, together with the complementary notions of ellipsis in Chinese and explicitation in the translation of Chinese into English. This is followed by a discussion of how attitudes, theories and practices may affect translation in language teaching and testing.

### **3.1 Defining translation**

My view of translation is not that it is a learned matching operation, but a profound and complex cognitive activity which consists of the following processes:

1. reading the source text in the source language;
2. comprehending the meaning of the source text;
3. relaying the meaning or message of the source text in the target language.

This three stage process assumes a knowledge of both source and target language and a creative cognitive process of understanding which enables transfer from one language to the other. Theories of the way in which this is carried out vary considerably, as described below throughout Sections 3.4 and 3.5.



### 3.2 Theories of translation

Theories of translation have paralleled reading theories to some extent in that there have always been 'top down' approaches, often labelled 'free' or 'loose', and 'bottom up' approaches, often labelled 'literal' or 'close'. The translation debate has probably not been so open to public view as the reading debate, for whereas literacy is now seen as central to the educational future of every human being, translation is not central to monolingual British life and is even now regarded as a somewhat elitist, intellectual practice. Baker (2001) has spoken of negative, even suspicious attitudes to translation in the UK. Nevertheless it is routine and essential to many communities, not only in the UK, but in bilingual communities throughout the world, bilingualism being the norm, rather than the exception in most linguistic communities.<sup>12</sup>

#### 3.2.1 'Bottom up' and 'top down' approaches to the theory of translation; prescriptive and descriptive views

It is necessary here to take two views of translation: how it should be done and how it is done. Traditionally, theories of translation have focused on the role of the translator as conveyor of a message and on the techniques used to convey that message to the target audience. In the following sections I will discuss the relationship of these theories to the models of reading described in Chapter Two.

By and large theories of translation have been prescriptive and have been concerned with two major considerations - fidelity to the message and elegance of expression. Some writers would go so far as to say that translation is in fact impossible - that there is no direct equivalent in meaning

---

<sup>12</sup> At the time of writing, there is considerable concern in the UK about the inadequacy of translation and interpreting facilities for the wave of asylum seekers from Eastern Europe and the Far East. Simultaneously, concern is being expressed about the lack of support for language learning at every level of the education system, viz. HEFCE Report on Chinese (1998), Nuffield Report (1999) etc.

or form in any two languages. But this is a somewhat impractical view, and translation may be approached from two standpoints:

a. The translator must adhere as closely as possible to the source text, so that the culture and the language of the source text are represented as faithfully as possible to the target language reader. This approach goes by various names - 'literal,' 'faithful,' 'close,' 'sourcier,' 'simpatico,' 'foreignised' and so on. From this point I will use the term 'sourcier' (pertaining to the source text and culture) as it is an appropriate and relatively objective term. The notion of equivalence is crucial to this approach, in which words or phrases in the source language will be reckoned to have optimal equivalents in the target language.

b. The translator should translate the message of the source text into a message that the reader of the target culture, the target audience, will understand in terms of his or her own language and culture. This approach is known as 'free' or 'cibliste,' may sometimes be regarded as paraphrase, and in certain cultural contexts is known as 'transcreation.' This approach I will refer to as 'cibliste' (pertaining to the target text and culture) the companion term to 'sourcier' and equally appropriate and objective. Naturally there are problems and advantages with both the sourcier and cibliste approaches and there are some writers who feel that some integration and balance between these two extremes is necessary.

### **3.2.2 The sourcier (bottom up) approach to translation**

From a cognitive point of view the sourcier approach may be likened in some ways to the 'bottom up' model of reading, progressing word by word through the text.

Professional translators may want to keep the 'feel' of a source text, and 'foreignise' their target text (that is, keep it as much like the original as

possible) sometimes even at the cost of 'disrupting the cultural codes that prevail in the target language' (Venuti 1995: 20). On the other hand some professional translators may prefer to 'domesticise' the translation, thus 'reducing it, if not simply excluding the very difference that translation is called on to convey' (Venuti 1995: 21). Venuti, discussing the relative merits of foreignisation and domesticisation, and taking Schleiermacher as his example, shows that 'to imitate .... (German) this closely is not to be more faithful to it, but to be more English, that is, consistent with an English syntactical inversion that is now archaic' (Venuti 1995: 102). Venuti himself abandoned the notion of *simpatico* because he 'became aware that...*simpatico*..mystifies what happens in the translation process' (Venuti 1995: 276). Foreignisation, or a *sourcier* approach, may have different effects on the reader; the prose may seem awkward and stiff, but on the other hand, rendering of culturally specific items may seem easier. For example, twenty first century British English readers would probably accept transliterated 'wok' more readily than they would accept domesticised 'frying pan' or rather un-English sounding 'cooking pot.'

Toury (1995) suggests that it is mainly European scholars who have 'had conceptual problems with target oriented work,' in other words those whose cultures had traditions of long, close contact, with considerable approximation of systems (Toury 1995: 175). Given the high incidence of cognates in, for example, Romance languages, the gap between foreignisation and domesticisation may not be so great when translating certain European languages. A modern exemplar of this attitude is Newmark (1991), who states that 'the literal word for word translation is not only the best, it is the only valued method of translation' (Newmark 1991: 10). Writing of inaccuracy, Newmark claims that 'literal translation can show up the depths of one nation's culture by converting it into words where emotions that have no target language connotations are exposed in their absurdity and

grotesqueness' (Newmark 1991: 44), surely an unwitting argument for less literal translation. Newmark claims to admire Halliday, who says that equivalence is lost at the level of units lower than the sentence (Halliday 1967). Halliday was a sinologist and knew full well that, as will be shown later in this chapter, there is simply no room for 'simpatico' in translation between languages as diverse in typology as English and Chinese.

### **3.2.2.1 The notion of equivalence in sourcier approaches**

The problem of equivalence has been widely discussed in the literature on translation. Catford (1965) claims that in 'normal' translation equivalence can occur between sentences, clauses, groups, words, and even morphemes. His examples, however, are French/English, and it is problematic to attempt to apply his claim to languages of widely differing typologies. In order to deal with cases in which equal rank equivalence was not possible between source and target text, Catford proposed rank shift and category shift. However, these deal only superficially with structures, and Catford admits that they involve departures from formal correspondence (Catford, 1965 :76).

The notion of equivalence is now widely debated, and in some cases rejected, by writers on translation. Snell-Hornby (1988), for example, remarks that 'equivalence...apart from being imprecise and ill-defined.....presents an illusion of symmetry between languages which hardly exists beyond the level of vague approximations...' (Snell-Hornby 1988: 22). As Jin (1995) points out, formal or semantic similarities are no guarantee of proximity in meaning (Jin, 1995: 231).

### **3.2.3 The cibliste (top down) approach to translation**

The cibliste approach may be likened in some ways to the top down model of reading, in which the reader/translator is guided by a holistic view of the text structure and by background knowledge and schemata.

It is important to recognise that translation is not just a matter of making linguistic choices. Lederer (1973) reminds us that 'le sens est donc la rencontre dans l'esprit de la formulation linguistique qu'on voit sur le papier et des connaissances dont on dispose à la lecture' (Lederer 1973:22). This encapsulates a rather more top down view of translation - text on paper, linguistic forms, reading and mind, all adding up to meaning.

Valéry (1936) points out that

*'il n'y a pas de vrai sens d'un texte. Pas d'autorité de l'auteur. Quoi qu'il ai voulu écrire, il a écrit ce qu'il a écrit. Une fois publié, un texte est comme un appareil dont chacun peut se servir à sa guise at selon ses moyens: il n'est pas sur que le constructeur en use mieux qu'un autre' (Valéry Variétés III 1936).*

Lederer's view and Valéry's view (quoted by Lederer) are consistent with the student centred, individualist approaches of the 1970s and certainly can be equated with Goodman's (1967) model of reading. Recognising both the author's intention and the reader's manipulation, Lederer makes a justifiable criticism of the sourcier approach:

*'si on traduit un texte phrase par phrase en s'inspirant plus de la langue originale que du continuum de la pensée de l'écrivain, on juxtapose des éléments linguistiques isolés qui correspondent individuellement d'une langue à l'autre, mais qui, assemblés, représentent un puzzle malajusté à la forme naturelle que prendrait la pensée dans l'autre langue' (Lederer 1973: 24).*

Lederer is convinced of the cognitive nature of translation and the importance of looking beyond the 'chain of signifiers:' 'Rejeter le sens en traduction, comme ont pu faire en linguistique les Bloomfield ou les Harris, ce serait rejeter la parole, c'est -à-dire le discours et sa finalité' (Lederer 1973: 26).

### 3.3 Descriptive views of translation

Prescriptive approaches to translation are, of course concerned with translator education - professional translators are often of the view that one should not translate until one has a perfect grasp of both the source and the target

language (Seleskovitch 1980: 128). Descriptive views are concerned with discovering what actually happens and using this knowledge to improve upon professional and pedagogical practice.

### **3.4 Recognising cognitive processes in translation - la saisie du sens hors langue**

A number of writers recognise that translation is more than matching a series of words in one language with 'equivalents' in another language. Lederer, for example, recognises the pedagogic value of translation as a means of exploring the target language (Lederer 1976: 69). Lederer's remarks, quoted above, show that the sourcier approach to translation does not allow scope for the cognitive aspects of translation which dovetail with the cognitive activity involved in reading. Seleskovitch (1976) emphasises the cognitive stage in translation - 'la saisie du sens hors langue' which must bridge the source text and the re-expression as target text (Seleskovitch 1976: 73). This is a simplification of what may be a very complex series or range of cognitive processes. To support her view, Seleskovitch cites Freud's preferred method of translation, which was to 'read a passage, close the book, and consider how a German writer would have clothed the same thoughts' (Jones, cited in Seleskovitch 1976: 84). She goes on to ask what happened once Freud had closed his book. He would inevitably have indulged in the restructuring processes common to all readers and translators, a process sometimes known as transcreation (Muhawi 1998).

Lefevere and Bassnett (1990) allude rather scathingly to the '*tertium comparationis*, the something which presumably hovers somewhere between languages in some kind of air bubble' (Lefevere and Bassnett 1990: 3), but which presumably is the translator's cognition, or 'la saisie du sens hors langue' suggested by Seleskovitch, above. Hartmann (1995), while espousing the notion of equivalence, and reducing translation to three stages - decoding,



recoding and encoding - still leaves room for the cognitive stage of recoding. He notes that of the three stages recoding is the most difficult, 'the crossing point at the border of languages' (Hartmann 1995: 514). Again, it is tempting to equate Hartmann's notion of recoding with *tertium comparationis* and with Seleskovitch's 'saisie du sens hors langue.' Most translators now would admit to some 'construction of meaning' in the process of translation (Richardson 1998: 125).

### **3.4.1 Means of providing evidence of cognitive processes in translation**

In recent years various experimental and descriptive methods have been used to determine the cognitive processes of translation, including think-aloud protocols, both simultaneous (for example, Harris reported in Harris 1992) and retrospective (for example Seleskovitch 1973), diachronic analysis of individual translators' manuscripts and synchronic analysis of parallel translations (for example Toury 1995). In addition, Harris has undertaken developmental studies of natural translation in bilingual children.

#### **3.4.1.1 Translation errors**

Hartmann looks for specific practical factors in translation, such as the differences between source and target language: differences in anaphoric systems; stylistic conventions and discourse signals; register and genre; cohesion; implicitness and explicitness. He concludes that it is not only interlingual difference which gives rise to errors, but that

*'errors are indirect evidence of, and unavoidable stepping stones on the way individuals approach the target. The study of these positive transfer and negative interference strategies ..... is therefore likely to be more useful than an abstract classification of micro-linguistic differences' (Hartmann 1995: 511).*

What Hartmann regards as errors, I believe can be accommodated on a cline of normalcy together with correct solutions, in line with Toury's theory of



norms in translation, and may be applied to a scalar or banded system of marking, discussed below in Section 3.6.3.4.

#### **3.4.1.2 Innate or natural translation**

Harris (1977) devised a 'manifesto' of natural translation, that is, the untutored translation carried out by bilinguals, as follows:

1. All bilinguals can translate and their competence is triple: competence in Li, competence in Lj and competence in translating both ways.
2. Translatologists should study natural translation, that is, translation done by bilinguals in everyday circumstances.
3. In natural translation, transmission of information is the prime aim and criterion of success (Harris 1977a ).

#### **3.4.1.3 Investigation of children's natural translation as a developmental phenomenon**

Harris studied bilingual children ranging from infancy to 18 years. He found that translation starts very early: in his data, the earliest recorded was at the age of 1;9; he found that spontaneous translation prevails in young bilinguals; that socially redundant and ludic translation is prevalent in young bilinguals; that children translate when they have not been encouraged to do so and even when they have been positively discouraged from doing so.

These findings led Harris to the conclusion that natural translation is an innate skill, in the weak sense of innate - a specialised predisposition in children to learn how to speak from the language they hear in their environment (Harris 1977b: 168). To support this view he cites the obvious pleasure children have in translating; their facility for associative memory for trans-word coding; their ability to recall content, irrespective of the language

of the original utterance; and in COMAL (conservation of meaning across languages or a transformation of symbols that preserves information)<sup>13</sup>.

Some scholars have criticised Harris on the grounds that his claims are based on very little data (Toury 1986), but the numbers of subjects, and the numbers of instantiations in his studies do not differ appreciably from those of many language acquisition studies.

#### **3.4.1.4 Think-aloud protocols**

Considerable value has been attached to protocol analysis as an investigative tool for the processes of reading and translation. Hartmann's view that think-aloud records and diaries provide the means of charting mental processes not normally accessible to direct empirical observation is representative of many (Hartmann 1995: 512). Lam (1995) claims that think-aloud protocols 'provide a valuable source of data about the sequence of events that occur' in translation and 'offer an immediate ongoing account' of internal language processing (Lam 1995: 905). I would like to suggest that the account given by the subjects is not as immediate and objective as she claims. Simultaneous think-aloud protocols are in fact an additional interpretation - the reader/translator arrives at his or her solutions, then proceeds to justify or explain them. Knowing that they are doing this for the benefit of a researcher may produce a halo effect and cause subjects to do some deliberate restructuring.

Krings (1987) criticises think-aloud protocols for four reasons: firstly, if the cognitive processes of translation are unconscious, the subjects are unlikely to have conscious access to them; secondly, human subjects are not unknown to produce verbalisations inconsistent with their actual behaviour; thirdly, the

---

<sup>13</sup> In his studies of children, Harris uses the term translation to refer to spoken utterances, often quite short; some of his subjects are too young to read and write. Researchers of adult translation usually

verbalisation task may alter during its performance; finally, such verbalisations may be incomplete reflections of the behaviour they are intended to describe (Krings 1987).

Faerch and Kasper (1986) note that often 'in order to meet the experimenters' expectations, co-operative subjects may verbalize more or even different information than they actually process during task completion....' and also note that on the other hand, verbal protocols may be incomplete (Faerch and Kasper 1986: 215).

In addition, introspection of this kind may not come easily to some subjects. Lam (1995) notes that 'most' of her subjects were able to think aloud after some 'practice.' This implies that some subjects simply could not think aloud, and also, if subjects need practice in this kind of task, it could well be that the behaviour as observed in the experiment is that learned from the researcher rather than the subjects' own spontaneous behaviour.

Even if readers are simply describing the processes that they sincerely believe have taken place there is already an added stage of cognitive restructuring. If the protocol takes the form of a retrospective diary, the time lag will involve an even greater degree of restructuring. I believe protocol analysis of translation runs the risk of being subjective and indirect.

Writing is a physical activity which records the process of a cognitive event or activity. On this premise, I suggest that written translation and its counterpart oral interpretation, whilst both still inevitably indirect, are less indirect protocols of cognitive activity than any additional simultaneous or retrospective introspection could be. Translation is, in effect, a 'write-it-down' protocol of reading.

#### **3.4.1.5 Diachronic individual analysis**

Comparisons of translations over time are rarely systematic and tend to follow the fashions of publishing and the exigencies of school curricula. The arguments between translators and reviewers tend to be heated and are rarely objective, as for example the debate over the many translations of Pushkin's *Onegin* (Nabokov 1955). Terms such as 'grotesque travesties,' 'dreadful verse,' 'teeming with mistranslations' are so emotional and subjective that they are unlikely to help translators, language learners or target language readers.

#### **3.4.1.6 Translation corpora**

The recent development of corpus linguistics has introduced a more systematic and objective way of analysing translation. The Translational English Corpus (2000), for example, is a computerised corpus of texts in many different source languages translated into English. It is possible to quantify characteristics of individual translators, and Baker (2001) claims that cognitive processes are visible in, for example, the insertion of unnecessary function words.

#### **3.4.1.7 Synchronic analysis**

Studies of the translations of groups of people operating on the same text at the same time are likely to reveal a great deal about the processes of translation.

##### **3.4.1.7.1 Translation norms: a descriptive account of synchronic parallel translations**

Toury (1995) has carried out a systematic descriptive treatment of translation theory, but still applies it to professional translators. Toury observes that translation processes do not lend themselves to direct observation, but he rejects introspection and comparison of translations made at different times (Toury 1995: 181).

Looking at 12 versions of one source text, Toury seeks a synchronic paradigm rather than a diachronic paradigm (Toury 1995: 188). He takes the view that what most often happens in translation is the opposite to Freud's method (Section 3.4 above) in that it is normally target text segments which are mapped onto source text segments, rather than the text as an entity. He points out that this segmental mapping can shed light on problems and solutions and in order to do this establishes the notion of translation norms (Toury 1995: 37). No doubt because he is dealing with professional translation, Toury bases his norms on established professional notions of adequacy with regard to the source text and acceptability in the target culture, and seeks a point between 'absolute rules on the one hand and pure idiosyncrasies on the other' (Toury 1995: 54).

He examines several parallel translations 'with an eye to drawing a distinction between what is universal and what is culture- or language-specific' (Toury 1995: 73). He does not specify cognitive distinctions, but his study strongly implies the inclusion of cognitive factors. He compares, for example, the translation of a metaphor as a similar metaphor, as a different metaphor, as non-metaphor and as zero, somewhat reminiscent of Catford's 'shifts' (Toury 1995: 82).

Toury's motivation is pedagogical and positive. Remarking that 'shifts' have been viewed as negative, he suggests that their identification is 'part of the discovery procedures only, ie a step towards the formulation of an explanatory hypothesis' (Toury 1995: 85). Toury's conclusion is that translating, even when done by children, is communicative and interactional in nature. The feedback received by the translator is normative, and emerging translators 'undergo a socialisation' (Toury 1995: 250). Toury sees the development of the native translator as a combination of nature and nurture, during which norms are internalised. (He further suggests that the

training of translators should encourage the violation of norms to bring about awareness of difference and diversity in texts and styles, but this is not central to my argument here.)

#### **3.4.1.7.2 Hofstadter's parallel translations of Clément Marot (1496-1544)**

A well known modern set of synchronic parallel translations illustrates the notions of norm and variation outlined above. In 1987 Douglas Hofstadter asked hundreds of friends and contacts around the world to translate Clément Marot's poem 'A une damoysele malade' (Marot 1537). His initial request asked translators to construct 'an artistic equivalent of Marot's poem in their own native languages.' In his request he noted that

*'for some people the meaning is sacrosanct; for others it is locally slippable but overall very important; and for yet others meaning at all levels and on all scales is something that can be freely toyed with. Different people, or the same person in different moods will make different decisions about what properties of the original are expendable and what ones must be preserved' (Hofstadter 1997: 7-8).*

This idea is not new, as can be seen from the account of translation theories given above. What is significant about the poem translations solicited by Hofstadter is that they give easily accessible published evidence of something like the norms described by Toury and Snell-Hornby.

Receiving hundreds of responses to his request, Hofstadter eventually reduced his 'thin wine' to a pure malt of 60+ versions which became the core of public lectures and a book. The book 'Le ton beau de Marot' consists of just over sixty (English) translations of 'A une damoysele malade' with lengthy, academic and often personal reviews and comments on each, including some acerbic comments on the three machine translations submitted. Hofstadter's contributors were all well educated adult academics, but the degree of spontaneity and infinite variety of their versions reflects what I have found in my data of teenage parallel translations.

Hofstadter does not make a statistical analysis of the variations in translation, but it is interesting to note that out of 62 versions, the phrase 'ma mignonne' (addressing the recipient of the poem) is translated in 62 different ways (let alone the rest of the poem). The most popular type among these includes some reference to 'sweet,' (11 instances). The second most popular type indicates smallness or youth as a term of endearment, such as 'babe', 'kiddo' etc. The third most popular type includes some reference to a sweet type of food, for example honey bun, pumpkin etc. as a metaphor for a loved one.

Hofstadter's contributors as a group show a cline of normative to non-normative readings, showing that a majority will go for some conventional, consensus reading, in a sense creating the norm, while a few read the source text in a different way, in a sense violating the norm. In any case, translation is highly individual, so that even within cultural norms, the translation choices reveal the schema of the individual reader/translator.

### **3.4.2 A note on automatisisation or proceduralisation in translation**

As with reading, a degree of automatisisation is present in the act of translating. The more experienced the translator, the more familiar the type of text, the greater the degree of automatisisation will be. But there will still be individual differences. Translation may to some extent be conditioned behaviour but every text is different and, as Toury says, 'a translation is something which hasn't been there before' (Toury 1995: 166). This observation echoes Bartlett's comment that memory has its own characteristics every time we make it, in other words that the differences are dependent, at least to some extent, on individuals' schemata.

In professional or expert translators there is available a conscious choice as to the degree of foreignisation or domesticisation - *sourcier* or *cibliste* - to apply to any particular text. But inexperienced intuitive translators will not



necessarily have this option. Whether their translation is *sourcier* or *cibliste* or balanced will depend on their knowledge of the two languages involved, their expertise in reading - and there is much to suggest in my data that some of the subjects may be fluent speakers but not good readers - and their background knowledge.

### 3.4.3 A note on cultural and chronological factors in translation

Snell-Hornby (1988) points out that the concept of culture is a totality of knowledge, proficiency and perception, and that it is connected with behaviour and events and dependent on norms. Translations will vary according to those norms (Snell-Hornby 1988: 40). This factor is of crucial importance in assessing the reading and translation of young people who will not have as complete a knowledge of 'culture' as mature adults. As will be discussed in Chapter Five, the development of schemata contributes greatly to general knowledge, and thus to reading and translation.

Similarly, time is a factor, highlighted by Lefevere:

*'Translations made at different times .... tend to be made under different conditions and to turn out differently not because they are good or bad but because they have been produced to satisfy different demands' (Lefevere 1990: 3).*

Venuti (1995) makes a complementary observation: 'when an author and a translator live in the same historical moment they are more likely to share a common sensibility, and this is highly desirable in a translation' (Venuti 1995: 273). The factor of shared or unshared time is probably at its most influential in young reader translators, either because they have as yet had no opportunity to experience, or because their peer group perceptions and attitudes cause them to see unshared time differently or even negatively. At sixteen years of age, readers may reject phraseology and ideology which they will accept at sixty. This will affect reading and translating quite profoundly.

Klein-Braley and Franklin (1998) cite two examples of problematic translation, in connection with unthinking use of dictionaries, which reveal the importance of general knowledge in translation. The first is the German word *Kulis* (Chinese labourers) translated as English 'biros' and the second is the 'howler' after which their paper is entitled, '*we have an alien on ice*' translated as something to do with foreigners in refrigerators (Klein-Braley and Franklin 1998: 60). Both examples reveal a lack of general knowledge, in the first example, about China, and in the second example, about sci-fi culture.

### 3.5 Translation into and from Chinese

Translation to and from Chinese is scantily known in the west, but gradually increasing. Westerners are usually aware of well known translations of a few landmark texts in Chinese literature and philosophy such as 'Monkey,' 'The Water Margin,' also called 'All men are brothers,' and 'The dream of the Red Chamber,' also called 'The story of the Stone,' and the major philosophers: Laozi, Zhuangzi, Mencius, Confucius and so on. Until the 1980s, when China opened to the West, the major effort in translation was largely unidirectional, that is from Western languages, chiefly English, into Chinese, and it is in this area that most of the theorising on translation has taken place. And while the Western reading public has embraced modern Chinese literature and film, particularly the dissident autobiography, with warmth, the balance is probably still the other way, as China looks to improve its science, technology and commerce. The problems are different from but complementary to those involved in translating from Chinese into English, and the crux of the task, the interpretation of grammatical and discoursal relationships between two languages totally different in typology, is common to translation in both directions.

### 3.5.1 Generalisations about the translatability of Chinese: ellipsis, implicitness, inexpressibility and ideation

As noted in Chapter Two and again in Chapter Four, there are certain characteristics of the Chinese language which affect reading and translation. These are mainly to do with absences of features which are present in European languages: inflectional morphology indicating person, number gender and case in nouns and adjectives; inflectional morphology indicating tense, person and number in verbs; definite and indefinite articles; inter-word spacing; overt pronouns; overt connectors, all of which are summarised in Section 2. 1.

#### 3.5.1.1 Extra-textual meaning in Chinese text: ellipsis and implicitness

I use the word ellipsis as a general term as defined by Sells and Wasow (1999): 'a construction in which something seems to be missing, but can be understood from the text' (Sells and Wasow 1999).

A common notion in Chinese writing and translation is that of implicitness: *yan wai zhi yi* - 言外之意 the meaning outside the words. A parallel term is *wen wai zhi yan* 文外之言 - the words outside the text, first used in connection with Buddhist sutras during the Tang dynasty. One of the earliest mentions of the term is in connection with the Wei Mo sutra.<sup>14</sup> Since then both terms have frequently been used to describe Chinese text and discourse.

As described in Chapter Two, the information in a Chinese text is implied as often as it is overtly stated, owing to the abundance of empty words, null arguments and the paucity of peritextual information ( no upper case, spacing

<sup>14</sup> Buddhist scriptures were translated by what can only be described as a democratic process, by a chief translator, who read out the text in its original Sanskrit and explained it. The explanation was interpreted orally into Chinese and written down by a scribe. Members of the audience, sometimes numbering several thousands, contributed their questions and suggestions. These were noted down, considered and incorporated into the final version which might take many years to complete (Qian no date:20-21).

between morphemes etc., as described in detail in Chapter Two). This is what in English would be termed ellipsis, but there is necessarily overlap between what is omitted in grammatical terms and what is hinted at in discursal terms. Moreover, omission is so common in Chinese that the term 'ellipsis' hardly covers the scope of the notion (see Sections 4.2.1 and 4.3.2 on redundancy, verblessness, deletion and the null argument in Chinese). Perhaps it is preferable to use a term 'close' to the Chinese, such as 'extra-textual meaning,' as what does not appear in a Chinese text includes the strictly grammatical as well as the discursal. Translation from Chinese into some other language, or translation from classical Chinese into modern Chinese requires the translator to 'fill in many blanks.'

### 3.5.1.2 Extra-textual meaning in Chinese text: inexpressibility

Another parallel notion to that of *yan wai zhi yi* is that of inexpressibility - *fei bi mo suo neng xing rong* - 非笔墨所能形容 (that which cannot be described with pen and ink). This concept achieves its most notable expression in works of Taoist philosophy, written in highly economical classical Chinese, for example: *dao ke dao fei chang dao* 道可道非常道 (the way which can be talked about is not the constant way). This quotation characterises the extra-textual meaning of Chinese and its stunning simplicity is only possible because of the absence of articles, the absence of a copular, the absence of verb inflection, the absence of noun inflection, the dual role of a single morpheme as either verb or noun and the flexible transitivity of the verb. (*dao* 道 may mean to talk about or to be talked about). Some analysts would regard this short text as two clauses, in which case there is an absence of correlative connectors ('if,' 'but,' 'and' etc.).

These characteristics of Chinese - ellipsis, implicitness, and inexpressibility - have no doubt been exploited by the elite literati in order to retain the mystery of education and learning, but they have filtered down throughout

Chinese culture and ideology. Huang Xuanfan (1977), refuting the idea that some things cannot be expressed in words, notes that Chinese philosophy has always centred on 'direct feeling' and is also 'wordless' (Huang 1977: 19). He points out that like any other language, Chinese has the capability of expressing in words whatever one wishes, as explicitly as one wishes. However, he feels that the traditional attitude has permeated Chinese language so as to become almost a habit, especially in the expression of first person feelings. He says:

*'It is not a coincidence that 'direct feeling' and 'wordlessness' have become the joint defining characteristics of Chinese philosophy and ways of thinking. Direct feeling is a process arrived at without having clearly undergone a reasoning out of the concept or without the support of evidence. In other words, awareness which does not make use of written or spoken language is 'direct feeling' (Huang 1977: 19).*

Huang's view of translation processes echoes that of Western cibliste writers in that he claims that regarding the process of translation simply as a transfer of textual structure 'does not accurately depict the psycholinguistic processes experienced by the translator' (Huang 77: 22).

Huang supports his argument by two observations: firstly, that Chinese is based on a topic comment structure (see Section 4.2.2.1), and, secondly, that the more words you have in a text the more the meaning comes through (Huang 1977: 40).

The traditional attitude outlined by Huang is exemplified in the research of Young (1994), who, invoking the notion of *yan wai zhi yi*, claims that the lack of inflection in Chinese makes it ambiguous, impersonal, and universal, and that the feature of zero pronouns and conjunctions makes Chinese poetry timeless and egoless (Young 1994: 97). Huang's scepticism towards this kind of attitude is supported by Cayley (1995), who questions this notion of universalization, pointing out that absence of pronouns in Chinese is relative,

and a matter of choice; in poetry a zero pronoun may simply be the result of a combination of good grammar and economy of expression (Cayley 1995: 767).

### **3.5.1.3 Extra-textual meaning in Chinese text: direct ideation**

Consideration of these characteristics of Chinese - ellipsis, implicitness and inexpressibility (manifest to some extent in all languages, but prevalent in Chinese) leads us back to the concept, adumbrated in Chapter Two, that Chinese characters provide a direct link to thought. Shen Xiaolong (1988) explains at length the view held by many Chinese and some Western thinkers. He claims that the pictorial element in Chinese characters transcends the constraints of time and space and the constraints of dialect and geographical boundaries. Without exploring factors such as practice, automatised and memorisation, he claims that modern psychology has proved that when reading phonemic orthography one can only understand the meaning by means of analysing the sounds, but that through square format Chinese characters one can access the semantic information directly through the picture (Shen 1988: 15). Shen is convinced that the principle of imagery on which Chinese writing is based increases the information content of individual characters, so that readers gain confirmatory information from connections in the context (Shen 1988: 13). This would seem to be contradictory - one might expect a higher information content in characters to lead to less dependence on context when reading.

It should be noted that Shen stands out as a champion of the Chinese way versus the 'great edifice of westernisation (which was) congenitally deficient' when it came to describing the Chinese language (Shen 1988: 5).

The view that Chinese characters link more directly to thought than alphabetic words, prevalent among well educated but perhaps linguistically naive Chinese and Westerners, may also contribute to the educational



emphasis on memorisation and received knowledge. It also bolsters the traditional prioritisation of 'full' words over 'empty' words, covered in Section 2.5.7.1.

Shen's concomitant views on combinations of characters in a text invite consideration. He claims that because Chinese is a non-morphological language, lexical, syntactic and semantic information is for the most part hidden in the linear flow of the arrangement of words (Shen 1977: 13). He supports this claim by quoting Humboldt's description of Chinese:

*'In Chinese sentences, each word is arranged so that one considers carefully each relationship before reading on in the text. Since the relationships of ideas are produced by these relationships in the text, this pure unadulterated silence takes the place of grammar.'*

*(Shen 1988, citing Humboldt, no date.)*

So, in writing on Chinese language and translation, on the one hand we have the notion that a character provides a direct link to thought and on the other hand we have the notion that a very great deal, that is, 'grammar,' is simply not expressed. In Chapter Two I adumbrated the notion that it may not be the orthography of Chinese which creates differences and or difficulties in reading, but rather the structure of the language, and in particular those areas where little information is provided grammatically, such as anaphoric reference. Shen Xiaolong's ideas might suggest that there are cognitive stages in reading that do not need to take place in the case of Chinese text. Are readers of Western languages wasting a huge amount of mental energy on verb, noun and adjective endings, hes and shes and ifs and buts? Or, conversely, are readers of Chinese wasting a huge amount of mental energy trying to imagine what the gaps mean? It is possible that analysis of parallel translations can provide clues to this dilemma.



### **3.5.2 The practical implications of ellipsis, implicitness, inexpressibility and ideation for the translator of Chinese**

The three properties of ellipsis, implicitness and inexpressibility pose problems for the translator concerned with rendering a Chinese text into a typologically different language which does not have these properties, particularly a European language. With regard to Chinese translated into English, the two languages focused on in this thesis, one of the main problem areas would be the interpretation of null arguments or zero pronouns, as appropriate noun phrases, pronouns or dummy subjects, or in certain cases as zero with an appropriate verb form. Ellipsis in Chinese usually, though not always, requires explicitation in English.

#### **3.5.2.1 Explicitation in translation**

Explicitation is the translator's practice of making information which is implicit in the source text explicit in the target text. It is the counterpart of implicitation, and according to some writers is synonymous with addition, while some writers regard addition as just one aspect of explicitation. Vinay and Darbelnet (1958), coiners of the term explicitation, say that it is information which is implicit in the source text, but 'can be derived from the text or the situation' which may be subject to explicitation (Vinay and Darbelnet 1958).

Nida lists situations where explicitation may take place, including elliptical expressions, grammatical restructuring, connectives, and categories of the receptor language which do not exist in the source language (Nida 1964: 227). Other writers believe that 'linguistic asymmetry' requires explicitation or, alternatively, implicitation, in translation (Klaudy 1996).

Explicitation may be optional or obligatory. Optional explicitation may take place where there are stylistic differences, or different text building strategies

across languages (Klaudy 1996). Explicitation is obligatory where syntactic and semantic structures, most notably the 'empty categories' differ in such a way that a translation would be ungrammatical without explicitation. The characteristics of Chinese outlined in Chapter Two and above in Section 3.5.1 usually necessitate explicitation when translated into European languages, particularly in the case of the null argument or zero pronoun.

Blum-Kulka (1986) claims that explicitation is inherent in the act of translation whether or not it is required by the text. She outlines the explicitation hypothesis, according to which there is 'an observed cohesive explicitness from SL to TL texts regardless of the increase traceable to differences between the two linguistic and textual systems involved' (Blum-Kulka 1986: 19). She suggests that empirical studies of professional and non-professional translation would show whether explicitation was a question of stylistic preference or an inherent trait. This may be shown by translators' means of showing cohesion - they may over-represent by lexical repetition instead of pronominalisation for example.

Blum-Kulka cites studies of language learners' translations, non-professional bilinguals' translations and professional translations, all of which show evidence of explicitation (Blum-kulka 1986).

Obligatory explicitation may be made more difficult for the translator, when, as in the case of Chinese, little grammatical information about person, gender and number is provided. Whilst obligatory, explicitation may not be straightforward.

In making explicit a zero pronoun, a reader/translator sets out in writing what he or she believes a proposition to be - providing an argumentless predicate with an argument, thus making it a full proposition. To do this

successfully, a reader/translator must have a very clear and confident understanding of the text. In the case of zero pronouns, patterns of explicitation of the null argument reveal the cognitive processes of the reader/translator. Explicitation of a zero pronoun as a pronoun or noun is an overt statement of a reader/translator's mental representation of certain aspects of the text, in particular, the arguments which constitute the chain of reference, and the propositions that make up the structure of the text.

### **3.5.3 Chinese approaches to translation**

The problems outlined above may be answered in part by looking at just a few examples of how in the past translators have tackled translation into and out of Chinese. This may provide some background for an analysis of my data from parallel translations by bilinguals.

#### **3.5.3.1 Fidelity, understanding and elegance: Yan Fu's late 19th century translations of Western literature and science**

Modern Chinese theories of translation began with Yan Fu (1854-1921) at the end of the 19th century. Yan Fu's work was a work of mission, to bring to a poverty stricken, backward nation facing colonisation at the end of the 19th century the technology and ideology of its aggressors to make it strong enough to match them. His prolific output included translations of Huxley, Dickens and Stuart Mill. Yan Fu is known for his three maxims of translation, the first being that of *xin* 信 (fidelity) but he has been criticised for being too free in his translation. It is likely, though, that a 'close' translation of, for example, Huxley, would not have made much sense in Chinese. His second maxim *da* 达 (to 'get it,' as one 'gets' a joke, ie understand) was more central to the task. *Da* is translated by some writers as 'intelligibility' or 'communicability.' His third maxim *ya*, 雅 (elegance) was aimed at making the work accessible to his readers. Yan Fu is sometimes acclaimed for having created literature out of non-literary texts and is given credit for giving the

Chinese literati access to Western thought. The reason for this is that at the time he was writing, *wen yan wen* 文言文 or literary Chinese (Mandarin) was still the norm in standard published Chinese texts. Although the vernacular language *baihua* 白话 had been used in certain written texts for some time, 'proper' writing was in the old literary style. Readers would have found a *baihua* translation of John Stuart Mill, for example, difficult to read mainly because of their expectations - a heavyweight scientific treatise would simply not be written in the vernacular.

On the other hand Yan Fu is sometimes criticised: for example, Huang Yushi (1995) alleges that 'great damage has been done by him to the general undertaking of translation in the country through a rather flippant statement made by him' that is, his maxims of *xin, da, ya*. (Huang 1995: 278). Huang implies that Yan Fu should not be regarded as the pioneering translator he in fact was. Huang's criticism is, I think, unfounded. Yan Fu's *cibliste* approach to translation made Western science and technology available to his target audience, educated Chinese. A *sourcier* approach would probably not have achieved this aim.

### 3.5.3.2 Lin Yutang - 'speaking for an old friend'

Lin Yutang (1895-1976), who, through his translations, defined Chinese philosophy for many Westerners in the middle part of the twentieth century, said:

*'Translation is a very subtle thing; unless you are emotionally in contact with the original author you cannot do a good job of it. In translating an author, you practically engage to speak for him in a new language, and you cannot do so unless you are speaking for an old friend, so to speak' (Lin 1961: 19).*

Lin is criticised by Huang Xuanfan (1977), however, for disregarding the cognitive factor in translation. Huang cites Lin's three maxims: the translator must have (1) a thorough understanding of the language and content of the

original text; (2) a reasonable standard of his own language so as to write comprehensible Chinese; (3) a proper view of the standards and techniques of translation. The maxims, however, are probably prescriptive maxims aimed at others, and can be assumed on Lin's part. The 'emotional contact' is mentioned as part of his own introspective view of his work.

### **3.5.4 Some more recent defining approaches to Chinese translation**

Since the days of Yan Fu Chinese translators have continued to discuss the principles and ideals of translation, and have continued to subscribe to the notions of *xin*, *da* and *ya*, though calling them by other names and each redefining them in their own idiosyncratic ways. More recent translators have had access to Western theories of translation and have been able to compare and contrast their work accordingly. Keenly aware of the mismatch in typology between the mainly European languages they translate from and into Chinese, they are more united in their approach than European translators.

#### **3.5.4.1 Huang Xuanfan - a *sourcier* to *cibliste* cline**

Huang Xuanfan (1977) makes a broad distinction between 'direct' translation (*zhi yi*) 直译 and 'meaning' translation (*yi yi*) 意译 which correspond to the western notions of *sourcier* and *cibliste*. He notes that such terms as these, and he includes Yan Fu's *xin*, *da*, *ya*, are fuzzy. His solution is to describe five points on a cline:

1. translating character by character and adhering to the sentence structure of the original;
2. translating sentence by sentence and adhering to the sentence structure of the original;
3. translating sentence by sentence, but not adhering to the sentence structure of the original;

4. translating according to the paragraph structure;
5. paraphrasing the gist of the whole text (Huang 1977).

Huang points out that the first option, though it would certainly fit in with some ideas of equivalence, as for example, outlined by Catford (1965), is patently ridiculous where Chinese is concerned. The last option is reminiscent of the notion of transcreation outlined by Muhawi (1998). Huang himself believes in 'meaningful' rather than 'direct' translation, inclining to the ciblister end of centre on his own continuum - that is free, but not too free, and manipulating the Chinese target text so that readers may understand the meaning of the source text. His is a balance between foreignising and domesticising a text.

Huang's position is that any theory of translation must be synonymous with a theory of semantics (*yu yi*) 语意. He addresses the subject of *yan wai zhi yi* - the 'meaning outside the words' which, as noted previously is of major significance in Chinese writing and reading:

*'Implicitness (yan wai zhi yi) is no mysterious thing. Everyday language has many ways of expressing implicitness. But that implicitness must be associated with the words, and must be comprehended by the listener through his own linguistic ability. The understanding of implicitness is through reasoning and it has its base in language' (Huang 1978:6).*

#### 3.5.4.2 Xu Suchong - expanding on *xin*, *da*, *ya*

Xu Suchong (1984) criticises extremes of formalism and freedom in translation and goes on to expand on Yan Fu's maxim. *Xin*, (faithfulness) becomes *zhongshi* 忠实 - also faithfulness, but with overtones of loyalty. *Da* (understanding) becomes *tongxun da yi* 通顺大意 - coherently accessing the meaning. *Ya*, he explains, is traditional or literary style.

Xu also adopts the broad distinctions of 'direct' translation and 'meaning' translation, defining them in his own terms. 'Direct' translation is putting

*zhongshi* ( being faithful and loyal ) to the content of the source text in first place, putting faithfulness to the form of the source text in second place and putting coherence of form third. 'Meaning' translation is putting faithfulness to the content first, coherence second and not sticking rigidly to the form of the source text (Xu 1984: 5).

Xu further delineates *zhong shi* (faithfulness) into the three aspects of content, form and style, each having three degrees of transparent, accurate, and correct in essence (Xu 1984: 23).

#### 3.5.4.3 Qian Gechuan - coherence and intelligibility

Qian (no date) adopts a cibliste approach: translation must be coherent and intelligible (*tongshun dayi*) 通顺大意 otherwise it becomes 'heavenly writing,' that is, no mortal can understand it. He borrows Matthew Arnold's suggestion that a reader should be able to forget that he is reading a translation.

Qian's reason for claiming that sourcier translation is impossible in the case of Chinese is the traditional notion that Chinese characters have evolved from images, while English is made up of letters representing sounds, as discussed above. Qian further claims that the logic of the two languages is different. Qian's examples of this difference in logic are, in the main, examples of different form rather than different logic, for example the difference between English 'there is/are' and Chinese *you* 有 . In grammatical context and in pragmatic use the two expressions are very similar.<sup>15</sup>

---

<sup>15</sup> The great 20th century Chinese grammarian Wang Li was also of the view that Chinese and Western minds think differently because of the difference in language. His key illustration is also the Chinese 'you' 有 - equivalent of 'there is.' However it is interesting to note that while Wang Li's views are supported by some Western scholars, such as La Polla, current Chinese thinking is that 'you' having a null argument, shows remarkable similarities both in form and function to structures in major European languages (eg: *hay, a, il y a etc.*) and differs from 'there is' in form only (Hong Kong City University Department of Languages and Translation Studies Colloquium, 11.12.2000).



Qian emphasises the role of reading in translation. Borrowing the poet Tao Yuanming's idea, he suggests that when simply reading, one reads for the *dayi* 大意 (gist) but when translating one must understand deeply and in detail.

What Qian cites as untranslatable *-zhe zi* 折字 or 'broken characters' in Chinese and palindromes in English - are linguistic puzzles and patterns. Their sole communicative purpose is to illustrate unique patterns in the language of which they are a part - there is no point in translating them.

### **3.6 Attitudes to translation**

Writing about translation has in the main been devoted to literary translation. In the UK, translation of scientific, technical and routine documents, along with spoken interpretation, has not been widely studied, perhaps because of elitist attitudes, or perhaps because of the world dominance of English as a second language. The area of translation studies has, however, developed in recent years and now encompasses studies of the more routine types of translation such as commercial, scientific and legal documents, especially in the European community, on an unprecedented scale.

#### **3.6.1 Attitudes to translation in language teaching**

Translation as part of language teaching methodology has had a chequered history. Language text books of the sixteenth century were often 'double manuals,' - parallel texts of English and French or English and Spanish, and so on - which depended entirely on translation. Describing Holyband's (1576) and Ascham's (1570) 'double translation' method, Howatt (1984) notes that in the hands of a skilful teacher

*'double translation ...gives equal status to both the foreign language text and the equivalent text in the mother tongue. It will not do to produce a hurried and half-hearted gloss if the ultimate aim is to recreate the original ..... accurately. The method is intended to make the learner equally conscious of*

*the structure and resources of his own language . Content is held constant while the resources of both languages are manipulated to express, as far as possible a common array of meaning (Howatt 1984: 34).*

Joseph Webbe (1560 - 1633), while pursuing the bilingual methodology, rejected translation on the grounds that construing in any language was impossible (Howatt 1984). Paul Festeau and Guy Miede were still using 'double grammars' (material presented in parallel in two languages) at the end of the seventeenth century, as was John Miller, in his book designed to teach English in Bengal (Howatt 1984: 53). The nineteenth century 'grammar-translation method' was an attempt to adapt methods used by adults learning by themselves (ie learning the grammar, then translating texts) for use in schools. Howatt points out that the main criticism of the method was aimed at the emphasis on sentences rather than whole texts. There was also a tendency, in Britain, at any rate, to make school translations unnecessarily difficult in order to prove that modern languages were not soft options, as compared with Greek and Latin (Howatt 1984: 135). Howatt points out that the disconnected sentences of the grammar translation method were no sillier than much of the 20th century audiolingual language teaching material (Howatt 1984: 141). Claude Marcel ( 1793 -1876) developed a very modern sounding 'rational' method, which included attention to the four 'branches' of language, that is, hearing, speaking, reading and writing, what we now term the four 'skills.' Marcel's method for teaching reading included a mother tongue translation which acted as an explanation of the target language text, and which enabled the learner to access a meaning ( in L1) directly.

A major change came in the late nineteenth century, when the reform movement, based on the primacy of spoken language and the science of phonetics, together with the new science of psychology, strove to banish translation from the language classroom. In the vanguard of the reform movement were Lambert Sauveur (1826 - 1907) and Maximilian Berlitz (1852

- 1921), who employed native speaker teachers (often more or less monolingual new immigrants). Their 'natural' and 'direct' methods allowed no translation under any circumstances. Henry Sweet (1845 - 1912), however, was aware of the importance of bilingualism in language teaching, believing that language teachers should be scientifically trained professionals who knew the phonetic system of both their own language and the language they taught. The five major areas of language learning contained in his 'Practical Study of Languages' are grammar, vocabulary, the study of texts, conversation and translation. Harold Palmer (1877 - 1949) while maintaining a monolingual approach, was not as extreme as the reformists, in that he acknowledged that withholding a translation might result in the student generating his own erroneous interpretation (Howatt 1984). While translation is not always discouraged in language teaching now (the 'Headway' English course books, for instance, provide occasional, tiny, translation exercises), it is not necessarily encouraged in language teaching. Malmkjaer (1998) notes that 'there are signs that translation is beginning to regain respectability among language teaching professionals - even in the EFL community where it has been particularly strongly villified' (Malmkjaer 1998: 1).

Some current writers concede the value of translation as an exercise at higher levels of language education: Newmark (1991), for example, regards it as the fifth skill, and 'important as an exercise in accuracy, economy and elegance' (Newmark 1991: 62), but this applies rather more to tertiary level than to secondary or primary level education. A more dynamic view is that of Hoenig and Kuszmaul (1982) who emphasise the communicative function of the target text, involving a shift in perspective on the part of the reader-translator who takes an active role in recreating the text (Hoenig and Kuszmaul 1982) a notion which reflects that of Marcel, above.

### 3.6.2 Translation in language teaching and language testing

As noted previously, some writers on translation believe that translation should not be practised until the target language is fully mastered - but when is that? Harris (1977), reviewed above, makes the point that bilingual children translate spontaneously and believes the phenomenon to be innate in the weak sense. Every language teacher knows well the habit of learners translating their innermost thoughts, sometimes awkwardly, into the target language. When we travel abroad to an area where the language is new to us, we decide what we want to say, consult a dictionary if necessary, then transfer it, the best we can, into the target language. Translation is intuitive and necessary and inevitable in the process of learning another language and also in the process of understanding another language through an intermediary.

It has been made unwelcome in classrooms in the UK for several reasons. Within the reformist movement of the late nineteenth century it was forbidden under any circumstances. While some language educators like Palmer and Sweet accepted its usefulness in the classroom, it has largely been displaced by techniques such as structural drills, role play dialogues and later by communicative activities (see Section 3.6.1. above). The grammar-translation method was associated with formal, literary language which was seen to have no place in modern conversation, and much of the reaction against translation was provoked by unreasonably difficult passages (See Section 3.6.1 above). Secondly, much of the methodology of language teaching in the second half of the twentieth century was based on, or at least inspired by the great 20th century wave of research into English language teaching. That wave involved native speaker teachers and, as many native speaker teachers of English as a foreign or second language are mono-lingual, translation would not be a practical possibility in many classrooms. Naturally, these teachers favour some kind of target language or direct

method teaching. Thirdly, in the beginning stages of language learning, direct method is highly beneficial, as every moment of classroom time is devoted to the target language and no time is wasted speaking L1. By contrast, translation inevitably means spending time speaking or writing L1.

### **3.6.2.1 Arguments against the use of translation in language teaching and testing**

Malmkjaer (1998) outlines the kind of objections to translation that have been made in recent decades:

1. It is different from and independent of the 'four skills' of listening, speaking, reading and writing;
2. It wastes time which could be spent on the target language;
3. It is unnatural;
4. It misleads students over the problem of equivalence, or one to one correspondence across languages and produces interference;
5. It is a bad test of language skills;
6. It is only suitable for training translators (Malmkjaer 1998: 6).

Malmkjaer allows that the 'doggedness' of these arguments suggests that they must be, to some extent, true, depending on the kind of experience students have of translation. My own view of these arguments is that they are dogma rather than dogged, and probably based on a somewhat unthinking faith in the post reform movement language teaching methods of the twentieth century. Firstly, it is scarcely necessary to point out that translation consists of reading and writing, interpreting consists of listening and speaking, and they are in fact perfect examples of 'integrated skills.' Secondly, the widespread use of translation, both professional and non-professional, would seem to contradict the view that it is unnatural. Thirdly, guided, tutored translation may serve to heighten students' linguistic awareness and prevent interference. Fourthly, investigation of translation as a test type is a relatively

new area of research compared with conventional modern language test types. Only after thorough analysis of translation as a test type will it be possible to make a justifiable judgement on its efficacy relative to other test types and in relation to the world of work and communication. Finally, to suggest that translation is only appropriate for training translators invites analogy with being taught to cook, or swim, and so on.

Objection 4, above, is probably due to the way in which translation has been handled - that is, with undue emphasis on the notion of equivalence, which, as discussed above, is outmoded.

There is an additional objection raised by opponents of translation in the classroom, and that is the problem of thinking. In Gatenby's words: 'We as teachers are trying to bring our pupils to use English without translating....Our aim is to get our pupils to the stage where they can use English without having to think' (Gatenby 1967). It would be fair to respond that thinking when using language is not a bad idea, and that thinking in one's own language does not necessarily prevent one thinking in another language.

Newson (1998) writing in the context of German students of English, in addition to using the type of objection cited by Malmkjaer, above, criticises the use of translation in teaching, as follows: it does not allow

- a) emphasis on initial fluency in spoken language;*
- b) ...controlled introduction of selected and graded structures (60's style) or communicative competence strategies (90's style);*
- c) ...controlled introduction of and mastery of selected and graded lexical items;*
- d) ...situationalised and contextualised language;*
- e) communicative language use*
- f) learner centred language learning*



*g) .....observable learning effect; (....since translation provides ...only one example of new language items; ....there is no repetition and practice..... no grading and no structuring (Newson 1998: 64).*

This unjustified, perhaps ill-informed attack seems to be assuming that translation is used to the exclusion of all else; in fact translation need only be a part of the broad spectrum of language activities. It also assumes random selection of texts for translation; in fact, translation texts may be selected in just the same, informed and purposeful way as any reading or listening text is selected for comprehension purposes, or in the same way that any communicative activity is devised for production purposes. When a whole text, rather than a single sentence, is set for translation, it provides its own context, and in addition, it can be talked through just as a reading comprehension may be talked through. Translation is communication, and is perhaps a more authentic form of communication than, for example, a contrived role play. Translation texts can be selected to reinforce structures learned in other stages of presentation and practice. In spite of his criticisms, Newson goes on to propose a model for teaching translation based on controlled use of a data bank of representative texts (Newson 1998).

Many of the objections to the use of translation exercises in the classroom are probably also based on experience of inadequate delivery. Any language exercise can be unsuccessful if it is delivered without due consideration and preparation. This holds true of testing also.

The current research does not seem to address certain perceived difficulties on the part of teachers which probably underlie anti-translation sentiments: the dominance of ELT methodology, often designed by and for monolingual English teachers; the difficulty of teaching translation effectively because of its known range of variation; the difficulty of marking translation, also



because of its range of variation. If teachers and students are accustomed to and trained in, the use of translation, many difficulties can be overcome.

### **3.6.2.2 Arguments for the use of translation in language teaching and testing**

I don't think that any language teacher would argue with the suggestion that target language teaching is undoubtedly the best way of teaching a language. But while there are distinct advantages in the use of target language only methods in the classroom, it is certain that learners instinctively want to, and do, translate on occasion, and there comes a point in every serious language learner's life when translation is necessary. It is one of the things we do with language and we should therefore channel or take advantage of this instinctive, or intuitive, behaviour to teach new foreign language behaviour.

Lu (2000) argues that Hong Kong learners of English depend heavily on their L1 and supports investigating how to make proper use of learners' L1 knowledge to enhance their L2 learning. Stibbard (1998) also working in the Hong Kong context, advocates a principled use of oral translation in the classroom, taking advantage of Hong Kong students' tendency to code-switch. Stibbard draws on Danchev's (1983) proposal that if students are going to use translation anyway, then it is better to acknowledge, channel and control it (Stibbard 1998: 70, Danchev 1983). Whether code-switching can be equated with translation is debatable, but is outside the scope of this thesis. Stibbard, like Newmark, suggests that translation is a fifth skill, alongside reading, writing, speaking and understanding.

Malmkjaer (1998) proposes five activities undertaken by a translator which are also essential language learning activities: anticipation; resource exploitation; co-operation; revision; translating. Malmkjaer further notes that to differentiate translation from the 'four skills' is misleading, in that it not

only necessitates competence in both languages involved, but also the additional ability to recognise the relationship between them (Malmkjaer 1998: 8). With regard to translation as a language test, Malmkjaer notes:

*'Since it is not possible to produce a good translation unless all of the traditionally recognised language skills have been employed during the process, an examination involving, say, the translating of a text for which sufficient groundwork has been done in class would, in fact, test the students' ability to apply all of the skills.'* (Malmkjaer 1998: 9)

Källkvist (1998) drew comparison between errors in compositions in L2 and translations from L1 to L2. She found that proportionately more errors occurred in translations than in compositions and that these included more meaning errors, but fewer collocation errors and over-use errors (Källkvist 1998: 85). She does acknowledge the effect of avoidance strategies in free composition, but takes no steps to control for this. A further shortcoming of her study is that the number of compositions analysed was only 8, as against 150 translations. It is interesting to note that she finds identifying errors in compositions more problematic than identifying errors in translations, a factor which should be taken into account in considering means of assessment. Källkvist's conclusion is that translation should be used in conjunction with other types of test (Källkvist 1998: 86).

An added advantage of translation is that, like any reading comprehension exercise, it opens up spheres of subject matter which might not be covered in strictly communicative activities. Language learners can understand more than they can produce, and can learn much from texts which are slightly beyond their production level. As pointed out by Qian (Section 3.5.4.3 above), translating requires reading in greater depth and detail, and is likely to have a greater impression on the learner than reading only. The additional cognitive effort required by the reader/translator to crystallise and express his or her comprehension of the text may contribute to learning. Translation does not need to be time consuming, but can be practised as part of an integrated

curriculum. As a written record of reading, translation can show exactly what the reader has and has not understood. The disadvantage of classroom translation is that it can only be effective if teachers are conversant with L1 and L2. It also requires preparation and guidance, and it may be difficult to mark. A possible marking approach for translation is discussed in Section 3.6.3.4.

Redvers-Jones (1995) notes that now, with the rise of 'post-communicative' approaches, language educators are beginning to see L1 as a resource for the learner rather than a handicap. Redvers-Jones lists three important roles for translation in language learning: firstly, the teaching of new items; secondly, as a way of learning and assessing both production and comprehension; thirdly, as a skill in its own right (Redvers-Jones 1995: 491).

Yu (2000) recommends the teaching of translation in Hong Kong secondary schools in order to improve confidence and creativity in writing, broaden students' cognitive and cultural perspectives, and raise students' awareness of the differences between L1 and L2 (Yu 2000).

Yan (2000) found that in teaching translation through the medium of L2 and L3, 90% of students found that not only their translation skills, but also their language skills improved. These students were Cantonese speakers learning English - putonghua (Modern Standard Chinese) translation and interpreting from a non-Cantonese speaking teacher. It would seem that the act of translation improves language ability.

### **3.6.2.3 Problems with question types in reading comprehension exercises - can translation help?**

An investigation of reading comprehension methodology is not central to this thesis, but in evaluating translation as a reading comprehension tool it is

necessary to review briefly some characteristics of reading comprehension exercises and problems associated with some elicitation techniques.

Widdowson (1978: 81), in criticising the use of excerpts for reading comprehension, suggests that initial paragraphs of a series could be presented in the mother tongue so that the student is, in effect, primed for the L2 reading task ahead and at the same time is provided with a whole text rather than just an excerpt. Presumably the teacher would need to translate the initial paragraphs first - an added burden some teachers might object to. However, Widdowson does at least admit the usefulness of mother tongue in the language classroom. Having criticised 'excerpts' from texts as inauthentic, he goes on to the problem of priming glossaries, which are, in effect, L2 paraphrases. His main point is that simplification may distort, and my view is that this problem would be overcome by straightforward translation - an activity which does not need to be done by the teacher alone - the whole class can contribute.

Widdowson's criticism is part of a sustained critique of reading comprehension methodology, targeting four common ways of eliciting students' understanding of a text, all of which, a quarter of a century later, are in common practice: wh-questions, polar questions, truth assessment and multiple choice (Widdowson 1978: 94). First of all, he highlights the oddity of asking questions about a text. In addition, wh-questions require no understanding of the text - the student simply needs to spot the key words in the appropriate syntactic structure. Polar questions, as testers well know, invite guessing, as there is a fifty percent chance of getting them right. Widdowson makes the point that with wh-questions and polar questions, if the teacher demands complete sentence answers, the student may be distracted because of the need to compose an answer in the target language, while with multiple choice questions, the distraction may lie in the need to

comprehend the alternative answers. Widdowson's conclusion is that truth assessment questions are the most effective form of questioning for reading comprehension, but the problem of guessing in relation to true/false questions is now well known. In evaluating reference questions in reading comprehension, Widdowson suggests that learners do not always have to understand what a sentence signifies - they simply need to understand the structure (Widdowson 1978: 101). The only reliable question type left, then, is the inference question, discussed in Section 1.4.8. Widdowson's example of inference is to do with anaphoric reference and requires very complex elicitation. Straightforward translation, as shown in my data in Chapter Five, provides clear evidence of understanding of anaphoric reference patterns. Target language comprehension questions do not always provide solid evidence of understanding and this is a gap which can in some circumstances be filled by translation into L1 or by summary in L1. It may be argued that translation is too difficult, but it can be graded and used with discretion.

### **3.6.3 Translation as a testing tool**

As noted above, it is not the main purpose of this thesis to evaluate language testing methodology or test types other than translation, but in discussing translation as a test type it is necessary to review some of the aims and requirements of language testing. Unlike almost any other test, a language test involves its object as its instrument, a situation which has interesting implications, if not problems. Modern test designs have moved away from this situation to a certain extent by including test types which use graphics or multiple choice of various types, especially where the aim is to test only reading or listening. However there are limitations on these test types. In testing language ability testers now look beyond form, taking a broad notion of communicative language ability or proficiency which includes the student's ability to understand discourse and have, for example, some sociolinguistic knowledge and also be able to deploy strategic thinking.

These types of knowledge are perhaps beyond the scope of graphics and multiple choice and more difficult to assess objectively.

### **3.6.3.1 The anomaly of measurement and authenticity**

It has been suggested by Bachmann (1990) that it may be possible to develop a precise definition of language ability that could serve as a base for developing a 'common metric' applicable to many different languages. The aim of this type of 'common metric' would be to provide reliable measures independent of specific languages (Bachmann 1990: 5,6). I am somewhat sceptical about this as my main area of concern is Chinese, the typology of which differs markedly from many languages, certainly from the major European languages on which most of the testing literature focuses, for example the Common European Framework for Learning, Teaching and Assessing Languages (1996).

One of the greatest problems in testing is marrying notions of language ability with principles of measurement. As noted above, we use language to measure language, so tests must be based on very clear definitions. I would like to suggest that in using translation as a test type we progress one small step in objectivity, in that we are measuring one language by means of another, the student's L1, and one small step towards a 'common metric' if that is possible, in that we are using cross-linguistic means.

Another dilemma in language testing is the use of authentic material - is there such a thing in language testing? If we espouse Widdowson's idea that interaction between the language user and the discourse is authentic (Widdowson 1978), then translation provides that kind of authenticity - not only in the task of understanding, but in concrete evidence of that understanding. Translation shows how the learner, in Oller's (1979) words,



understands 'the pragmatic interrelationship of linguistic context and extralinguistic contexts' (Oller 1979).

One of the characteristics of many current language tests is that they aim to test only one ability. But real life language requires simultaneous use of many abilities - 'talking' is sometimes regarded as a fifth, integrated skill, alongside listening, speaking, reading and writing - and Newmark (1990) has suggested that translation, the bilingual integrated skill, is the fifth skill. Bachmann points out that some tests which are designed to be unidimensional cannot be authentic, while tests of authentic language use may not provide reliable measurement (Bachmann 1990). Translation is a task which may, to some extent, overcome these anomalies.

### **3.6.3.2 Testing communicative competence by means of translation**

Translation is widely perceived as the domain of experts and virtually impossible to evaluate objectively. But it is also widely agreed that language tests are not the models of objectivity that some believe them to be, for every test is set by someone who believes his or her own selection of items and criteria to be right. The very objectivity of the marking procedure indicates a highly subjective manipulation of test items. Questions used in objective test types do not always give the learner scope to show all that he or she has understood, or to show how the answer has been arrived at. Candlin (1986) writes that communicative competence is

*'the ability to create meanings by exploring the potential inherent in any language for continual modification in response to change, negotiating the value of convention rather than conforming to established principle. In sum ... a coming together of organised knowledge structures with a set of procedures for adapting this knowledge to solve new problems of communication that do not have ready-made and tailored solutions' (Candlin 1986).*

This of course takes in all those little things which are difficult to test, and indeed difficult to validate empirically, such as sociocultural or sociosemantic



knowledge, cohesion and coherence, degrees of formality and so on. It is the interaction of these various aspects which characterises competence, and, with regard to reading comprehension, translation may provide a very appropriate means of testing that special kind of interaction.

### **3.6.3.3 Testing strategic competence by means of translation**

A major aspect of communicative competence which can be assessed by means of translation is strategic competence, not only when the learner's linguistic knowledge alone is insufficient to tackle the text, but also as a means of enhancement. It may be argued that strategic competence includes abilities which are not necessarily linguistic, but have to do with general intelligence - but is this not the case with most uses of language? Certainly in the case of reading comprehension, every reader must draw on resources of background knowledge to deal with new information and must make inferences from what is in the text in order to deal with what is implicit only. Translation exercises provide evidence of what strategies a reader has employed to work out meanings and relationships hitherto unknown. This is evident in vocabulary choices, where students' miscues reveal inappropriate use of schemata, sometimes due to phonological and visual misreadings. Student translations of verbs may reveal reversed transitivity (dog bites man/man bites dog) while student translations of subordinate clauses may reveal linear reading (ie where L1 is left-headed and L2 is right headed, students may read and translate from left to right, and thus mis-assign modifiers). The problems of propositional representation (who does what to whom) are clearly seen in the way students translate pronouns and other anaphors. Translation reveals all these problems without the need either for specially designed questions or for introspection.

Not only could translation be used in teaching reading comprehension, but in the same way could be used in testing comprehension as part of a test battery

or as a single test. The problem, as noted previously, is marking. Like continuous or creative writing, translation is difficult to mark because of the range of possible answers.

#### **3.6.3.4 Marking translation and criteria for correctness**

One of the main arguments against translation in language education is the difficulty of marking and evaluating. The advantage of modern, objective test types such as multiple choice, cloze, short answer questions etc., is that answers are right or wrong and can be marked and standardised easily. Use of these test types means that markers can be trained easily, marking can be done quickly, and statistics on results are clear and manageable. In a situation where league tables exercise enormous influence on schools, pupils, teachers and parents, it is important to have statistics which can be easily read and interpreted, and objective test types contribute greatly to this. Where testing is less objective, as in the case of creative writing and translation, the onus is on the marker. Markers require more thorough training, marking takes longer, marks may be contested by candidates and schools, and statistics on results may be less easy to compile, less clear, and less easy to interpret.

Bachmann shows how giving instructions to test takers about criteria for answers, for example, grammatical correctness and breadth of vocabulary, may affect the way in which they answer questions (Bachmann 1990) and also points out that there are measures which do not require a correct answer - such as questions on attitudes and feelings. Translation is an area which gives scope for a variety of answers, as long as test takers know beforehand what the criteria are. If they know that a translation task is an opportunity, for example, for them to show that they can deal strategically with language, and that there is scope for more than one type of answer, they may blossom.

The response of a test taker may not always be what either the setter or the marker expects. Bachmann notes that problems in scoring and interpreting unexpected responses lie in deciding whether such responses reflect the test taker's 'normal language processing' or incompetence (Bachmann 1990: 126). Bachmann further notes that reliability and validity may be compromised by the problem of the unexpected response. Crucial decisions on how to score, or whether to penalise, have to be made, but Bachmann notes that quantitative analyses of actual responses, including unexpected ones, may help to clarify this kind of situation. The analysis of vocabulary responses outlined below in Section 3.6.3.4.1 and the putative marking scheme developed from it may provide a solution to the problem of unexpected answers, one of the chief problems with translation. The responses obtained in this study, that is, interpretations of zero pronouns in Chinese texts, lie within a narrower scope of variation and may also provide a possible approach to marking and are explored fully in Chapter Five.

#### **3.6.3.4.1 A further exploration of marking translation**

While direct equivalents in translation are difficult to achieve, it is possible to arrive at a norm, and a cline of normalcy on which the norm or consensus solution may be viewed as correct, and other solutions may be regarded as nearly right, 'acceptable' and unacceptable. Studies by Toury (see Section 3.4.1.6.1), analysing student translators' parallel translations, and by Hofstadter (see Section 3.4.1.6.2), using adult experts' parallel translations, show similar results to those of my data, gathered from teenage intuitive translators' parallel translations.

When vocabulary is translated (within a text), groups of parallel translations show distinct clines (Waggon 1999a). A group of 100 reader/translators were required to translate a Chinese text which contained the item *zawen* - 'essay.' The most popular answer (28 out of 100), was 'essay' and coincides with what

would be regarded as the norm, or consensus reading. This may be termed the homoseme and on a marking scale can score full marks. 18 readers opted for 'writings' perhaps not knowing exactly what the word was, but able to understand from the context that it was some kind of writing, thus using a more general term. These answers may be termed parasemes, and score nearly full marks, or 'B.' Other renditions included a range of journalistic writing types eg 'journal', 'article', a range of literary writing types eg. 'short story,' 'novel,' 'poem,' and a range of pedagogic writing types eg. 'excerpt,' 'passage,' 'comprehension.' These may be given credit for working out the meaning from the context, but are less accurate than the more general 'writings.' These may be termed allosemes and would score satisfactory, or 'C' marks on a scale. This type of answer may show strategic competence. Finally there are those solutions, usually in the minority, which do not relate at all to the target text (for example 'dragons') and may be termed nonsemes and deemed worthy only of minimal or zero marks (Waggot 1999).

This type of categorisation would need to be combined with the conventional scalar approach to writing according to which texts are marked globally. If test setters and markers assume, and work with, the notion of interactive reading as described in Chapter One above, there is no risk of incompatibility, since the solutions to individual vocabulary items can be assessed as part of the reader's understanding of the text as a whole.

In Chapter Five this type of cline is adapted to describe parallel translations of reference chains, focusing on the zero pronoun in Chinese, and could be adapted to deal with morphological features, syntactic structures, and coherence in discourse.

### **3.6.3.4.2 Application of marking schemes for translation**

The marking scheme outlined above is putative and has not been tried out. It would be necessary to design appropriate marking schemes according to the requirements of particular tests and testing authorities. Careful, co-operative selection of source texts and rigorous training of markers would be crucial to efficient operation of such a scheme.

## **3.7 Summary and conclusion**

Studies of professional translation show that approaches differ widely, from *sourcier* to *cibliste*. This range of possibilities may make translation difficult to teach and to mark. The prescriptive attitude of professional translators is not encouraging in an educational context. However, studies of translation in children, studies of translation errors, studies of protocols, recent developments in translation corpora, and studies of parallel translations, show that translation may be approached descriptively.

Three major characteristics of Chinese text - ellipsis, implicitness and inexpressibility - indicate that translation of Chinese is necessarily *cibliste* in approach, as illustrated by the work of well known Chinese translators and writers on translation throughout the twentieth century (Lin 1961, Qian (1981), Huang 1977, Xu 1984, Huang 1995,). The 'empty categories' of Chinese necessitate explication (Blum-Kulka 1986) in English translation, which could reveal much about the cognitive processes of translation (Baker 2001).

Although throughout the twentieth century there has been some hostility towards translation in language teaching (Malmkjaer 1998), the history of language teaching shows that bilingual approaches have always been used and have not been unsuccessful (Howatt 1984). Some language teachers are now accepting that translation may be used in class in a controlled way. While there are objections to using translation in the classroom, some teachers

have found that it has advantages (Malmkjaer 1998, Kallquist 1998, Lu 2000, Yu 2000, Yan 2000, Stibbard 1998). It can, and perhaps should, be used for teaching and testing language, and it is possible to train teachers and testers to use appropriate selection of tests, appropriate teaching techniques and appropriate, objective marking schemes.

## **Chapter 4: The role of the null argument or zero pronoun in Chinese discourse**

### **4.0 Aim and introduction**

The aim of this chapter is to review research on anaphoric reference and in particular on the null argument in Chinese. A description of important characteristics of Chinese which affect comprehension of the null argument or zero pronoun in reading expands on the brief description in Chapter Two. These characteristics include the lack of inflectional morphology, the tendency to pro-drop and topic prominence. The topic chain is closely related to the pattern of anaphoric reference. Anaphoric reference is viewed from two points: anaphoric choice, which is the speaker's or writer's decision on whether to use a NP, a pronoun or a zero pronoun to refer to an antecedent, and anaphoric resolution, which is the listener or reader's decision on what is the antecedent of any particular anaphor. Patterns of anaphoric choice in languages other than Chinese, then in Chinese, are discussed from the point of view of syntactic structure, and from the point of view of discourse analysis. While anaphoric reference itself is a syntactic device, the type of anaphor used may be governed by criteria of discourse. The chapter concludes with a review of some of the few experimental studies of anaphoric resolution of the null argument in Chinese.

### **4.1 Null arguments, zero pronouns and anaphora across languages**

The phenomenon of the null argument or zero pronoun is well known to learners and teachers of Chinese and is increasingly well documented in the literature (Huang Yan 1994; C.-T. J. Huang 1989; Li and Thompson 1979; Tao 1996; Wong 1992 and others). English speaking learners of Chinese, for example, sometimes find it difficult to minimise the use of pronouns in Chinese when speaking and writing, and sometimes misinterpret null arguments, or zero pronouns when listening and reading as they are



accustomed to a language which maximises the use of the pronoun. There is also some evidence to suggest that native Chinese readers and listeners also sometimes misinterpret null arguments (Tao, 1996, Wong 1992, Young 1994), as do speakers of other languages where pronouns or null arguments may be ambiguous. Leong and Hsia (1996), in a paper criticising the borrowing of English word order and syntactic patterns into Cantonese, have documented a tendency among young Hong Kong speakers of Cantonese to insert 'a redundant subject when the verb object pairing is made clear by context' (Leong and Hsia 1996: 76). In other words, young Chinese people who have access to English are rejecting the null argument or zero pronoun in Chinese, perhaps because they find a more explicit pattern more convenient.

#### **4.1.1 A note on pro-drop languages**

Languages differ in their patterns of pro-drop. Herman and Yoon (1989) propose four types of pro-drop language: (i) core pro-drop languages which have obligatory expletive and optional referential pro-drop (eg Italian and Chinese; (ii) core non pro-drop languages which allow neither expletive nor referential pro-drop (eg English and French); (English does allow pro-drop in certain genres and certain circumstances eg diaries, headlines, etc. and in certain environments, for example in a series of verbs having the same argument.) (iii) restricted pro-drop languages which allow either expletive or referential pro-drop in certain contexts (Bavarian German and Old French); (iv) expletive pro-drop languages which allow expletive but not referential pro-drop (Hermon and Yoon 1989). These are, in the main, cases of languages which have rich inflection which provides information about number, person and gender, thus enabling the listener/reader to make accurate anaphoric resolution.

Chomsky (1981, 1982b) set up the pro-drop/null subject parameter, which has recoverability as a criterion for distribution of pro-drop, that is, a pronoun

may be dropped if its content can be recovered in one way or another. The condition for this is the existence of rich inflectional morphology which includes an elaborate system of agreement. The model predicts that an argument may only be encoded as zero if its verb agrees. This may generally apply to arguments in subject position, but not to arguments in object position, as typical Italian type pro-drop languages and typical English type non pro-drop languages do not have verb-object agreement. The importance of the verb in anaphoric choice and processing of null arguments in Chinese will be discussed in Section 4.2.2.2 below.

While many of the so-called pro-drop languages assist the speaker's use and the listener's comprehension of zero pronouns by means of the inflectional endings of verbs (eg *hablo español*) and inflectional endings of adjectives which indicate case, gender and number, English has relatively few of these helpful clues, and information for the listener or reader about gender, number and person is contained in the pronoun. Chinese however, while having abundant pro-drop has no useful inflections to help establish the substance behind the zero. While English verbal ellipsis is heavily dependent on operators, which give information about number, person and tense, Chinese must rely for ellipsis on re-iteration of the lexical verb. If the verb is the least dispensable element in a clause in English (de Beaugrande and Dressler 1981: 67) it is equally, if not even less so, in Chinese (Waggoner 1997a: 4). Even so, it is evident that Chinese also has verbless sentences, as described in Section 4.2 below. A parameter which relies on a rich agreement system clearly has no relevance for Chinese.

#### **4.1.2 Patterns of reference**

Brown and Yule (1983) maintain that

*'pronouns are the paradigm examples of expressions used by speakers to refer to given entities....(they) are phonologically and lexically 'attenuated.'*

*Because of their lack of content they have become the crucial test case items for any theory of reference' (Brown and Yule 1983: 214).*

Halliday and Hasan (1976) suggest 'a logical continuity from naming, through reference to textual reference' (Halliday and Hasan 1976: 32). In practical terms this may be realised as the following basic pattern:

N or NP (referring exophorically or endophorically) > PRONOUN (referring endophorically) > ZERO (referring endophorically)

Huang Yan (1994) regards this pattern as a semantic content hierarchy and notes that 'the inherent semantic content of a lexical NP tends to be semantically more specific than that of a pronoun, and the inherent semantic content of a pronoun, than that of a zero anaphor' (Huang 1994: 16).

Brown and Yule point out that pronouns are most often used in situations of anaphora, referring back to a given entity in the text; exceptions, where speakers may use a pronoun to indicate something or someone in the immediate environment, may be regarded as deixis (Brown and Yule 1983: 214, 215). They discuss briefly the notion that the antecedent predicate, rather than the antecedent noun phrase, may affect both speaker's and listener's interpretation of the pronoun, allowing the use of anaphoric patterns which are not strictly grammatical or logical, but perfectly comprehensible, as in :

21.

Even an apprentice can make over twenty pound a week and they  
don't get much tax [taken] from that. (Brown and Yule 1983: 217)

This notion reflects the issue, addressed in Section 2.6.3.1.1, that the verb in Chinese text assists the reader or listener in anaphoric resolution.

In contrast with the basic reference chain of English described above, a Chinese chain of reference normally begins with a noun phrase and continues

with zeros, and may even apparently begin with zero. This notion will be discussed more fully in Section 4.2.2. This happens in European languages, for example in Spanish, as noted above, and in Italian, where phonetically null subjects are allowed in tensed clauses, for example ‘verrà’ he will come’ (Aoun 1985: 50). But in these cases at least some information (though perhaps not complete information, for example number *or* person) about the argument associated with the verb is given in the verb ending. By contrast, not only are pronouns in Chinese ‘phonologically and lexically attenuated’ to the point of zero, but there is no ‘verb ending’ to give information.

## **4.2 Characteristics of Chinese grammar which affect choice and comprehension of null arguments**

### **4.2.1 Morphological structure of Chinese**

As described in Chapter Two the basic unit of Chinese may be said to be the morpheme which in most cases is written as a single character and in most cases is uttered as a single syllable. Chinese is regarded as an ‘isolating’ or ‘analytic’ language : it has no capability of inflection, that is, person, number, gender, case, tense, and aspect are not indicated by any morphological change to the syllable. Additional syllables are added to lexical items to indicate, for example, aspect, but in writing these are not ‘joined’ to the preceding syllable, owing to the practice of regular intercharacter spacing in Chinese. In the main, information on person, gender, number, case, tense and aspect is provided lexically rather than morphologically. Where knowledge about person, number, gender, case, tense and aspect is considered to be background, or given or shared, that information is frequently not overtly provided in the text. There are no definite or indefinite articles as such; information about definiteness is provided by, for example, using demonstratives.

4.2.1.1 A note on redundancy

This high degree of succinctness begs the question of redundancy, which in European languages is largely carried in morphological inflection. Redundancy in Chinese is mainly carried in the classifier system. Every noun has a classifier:

|     |                   |                    |                |
|-----|-------------------|--------------------|----------------|
| 22. | 一本书               | 这杯咖啡               | 哪个人            |
|     | yiben shu         | zheibei kafei      | neige ren ?    |
|     | one (volume) book | this cup coffee    | which (class.) |
|     |                   |                    | person?        |
|     | A book            | This cup of coffee | Which person?  |

In answers to questions and in anaphoric reference the noun may be omitted. Redundancy is also carried lexically, in that words are usually formed by combining two morphemes, or characters (see Section 2.1) and frequently these morphemes are of similar or related meaning. For example ANJING 安静 (quiet) consists of two morphemes, *peace* and *stillness*. Thus, the apparently extreme economy of Chinese does not reduce its redundancy.

4.2.1.2 A note on verblessness

Verblessness is not central to the argument of this thesis, but it is worth noting here that verbless sentences are a relatively common phenomenon and part of the pattern of economy in Chinese. Tang (2000) cites a number of examples, within a minimalist framework, most of which omit the copula:

23a.

- 张三中国人
- Zhangsan zhongguo ren
- Zhangsan China person
- Zhangsan is Chinese.

23b.

张三傻瓜

Zhangsan sha gua

Zhangsan silly melon

Zhangsan is a fool.

23c.

一杯咖啡五块钱

yi-bei kafei wukuai qian

one cup coffee five dollars

A cup of coffee costs/is five dollars.

Tang notes that the interpretation of the empty verb in verbless adverbial clauses is restricted to the verb to be, and that the interpretation of the empty verb in empty verb clauses depends on the discourse (Tang 2000).

#### 4.2.1.3 Deletion and transposition in Chinese

The widespread phenomena of ellipsis and implicitness, especially as observed by translators, were discussed above in Section 3.5.1. Lu (1986) maintains that implicitness is quite different from ellipsis or deletion: 'Strictly speaking, deletion can only be used where the item is a reinforcing item, otherwise, it must be regarded as implicitness' (Lu 1986). According to Lu the complementary feature of deletion, transposition, can only be carried out under certain conditions, and is subject to syntactic constraints (Lu 1986). Supporting Lu's argument, Xu (1991) points out that transposition leaves traces in the form of empty categories as in the following examples:

24a.

我相信张三喜欢李四

wo xiangxin Zhangsan xihuan Lisi.

I believe Zhangsan likes Lisi

I believe that Zhangsan likes Lisi.

24b.

张三，我相信喜欢李四

Zhangsan, wo xiangxin 0 xihuan Lisi.

Zhangxin, I believe (zero pronoun)likes Lisi

I believe Zhangsan likes Lisi.

24c.

李四，我相信张三喜欢

Lisi, wo xiangxin Zhangsan xihuan 0 .

Lisi, I believe Zhangsan likes ( zero pronoun)

I believe Zhangsan likes Lisi.

Xu goes on to say that any kind of transposition in a sentence is marked, and that transposition and deletion are interdependent. They produce flexibility, but are liable to principles and constraints (Xu 1991: 70).

Zhang (1993) reviewing Lu's theory on transposition and deletion, shows how it may be the part of speech which governs deletion, as in the following example:

25.

你一言我一语

Ni yi yan wo yi yu

you one word I one word

you say a bit, then I say a bit.

Zhang points out that one can only say for sure that it is a verb which is deleted, but it could be one of a number of verbs (eg shuo, lai etc.) (Zhang 1993: 126). This is an example of the verblessness described by Tang, reviewed in Section 4.2.1.2 above.



Zhang does not draw a hard demarcation between deletion and implication, saying that the elements of implicitness are made clear from the environment of the sentence, including the context, and can also be recovered from the sentence itself. The latter is particularly important, because it is within the sentence that rules operate (Zhang 1993: 126). According to Zhang deletion is similar, yet also dissimilar to implicitness: if a word or phrase departs from the context, or if the context is not clear, extra meaning must be added. Only certain specific elements can be deleted, and, once elements are added, the structure must be legal. The second part of the sentence makes clear what the deleted element is, and once it has been added, it is a legal sentence (Zhang 1993: 127).

#### **4.2.2 The notion of topic in discourse in languages other than Chinese**

The notion of topic in discourse is now widely recognised as a feature crucial to language production and comprehension and is crucial to patterns of reference. Givon (1983) designed a model of description for a unit he termed the thematic paragraph within which to describe continuity in discourse. He proposed three major aspects of continuity: thematic continuity, action continuity and topic/participant continuity.

*'The thematic paragraph is by definition about the same theme. Most commonly it also preserves topic and action continuity. However, topics/participants may change within the discourse without necessarily changing either action continuity or theme continuity. And action continuity may change without necessarily changing thematic continuity. One is perhaps justified in viewing the three as an implicational hierarchy (or 'inclusion set') (Givon, 1983: 8).*

This is a suggestion which allows for description and analysis of anaphoric choice and resolution in languages where strictly formal grammatical principles (ie inflectional morphology) may not apply. It could be applied to Chinese for example, where in some cases the verb must be relied upon as a

cue for the understanding of abundant zero anaphora (discussed later in Section 4.3.2.2).

The model may not be entirely suitable for Chinese, as Givon states that ‘In the ..... clause level coding (syntax) action continuity receives its expression strongly and universally via the tense-aspect-modality sub-system most commonly attached to the verbal word’ (Givon 1983: 8). The ‘clause level coding’ of syntax may be somewhat weaker in Chinese than in, for example, English. There is, for example, no sure way of distinguishing between finite and non-finite verbs in Chinese. Huang (1994) suggests that every Chinese verb is finite (in terms of Chomsky’s definition) (Huang 1994: 30). While aspect may be indicated in Chinese, tense is not marked. It may be that Givon’s three types of continuity may be indicated in other ways in Chinese than in inflected languages. The temporal sequentiality which Givon associates with action continuity within the thematic paragraph would in the main be indicated lexically in Chinese by time adverbs and by connectors such as JIU 就 (*then*).

Givon’s model is strongly linear. He proposes three types of topic: chain initial, characteristically newly introduced and discontinuous; chain medial, characteristically continuing and persistent, but not maximally persistent; and chain final, characteristically continuing but non-persistent (Givon 1983: 9). He proposes a continuum of availability to the hearer of high to low referential distance, consistent with low to high continuity, and high to low persistence, consistent with high to low continuity. Givon’s continuum makes clear the relationship between anaphoric reference and topic continuity.

Givon's continuum is as follows:

*most continuous/accessible topic*

zero anaphora

unstressed/bound pronouns or grammatical agreement

stressed/independent pronouns

R-dislocated DEF-NPs

neutral-ordered DEF-NPs

L-dislocated DEF-NPs

Y-moved Nps (contrastive topicalisation)

cleft/focus constructions

referential indefinite Nps

*most discontinuous/inaccessible topic*

(Givon 1983)

As noted above in Section 4.2.1 definiteness may not necessarily be overtly indicated in Chinese, so the three central points on Givon's continuum may not be directly applicable to topic continuity in Chinese. Shi (1989) points out that in Chinese while initial mention of a topic tends to be indefinite, subsequent mentions of a topic tend to be definite, but this definiteness is not overtly marked in the linguistic forms - it may be expressed as zero (Shi 1989).

According to Givon, referential distance can be measured in number of clauses to the left and persistence (or decay) can be measured in terms of number of clauses to the right (Givon 1983: 13-15). As will be seen later, a number of researchers (eg Fox 1987; Tao 1996) feel that a linear approach is inappropriate, notably with regard to the phenomenon of return pop, the term coined by Reichman (1981) for the re-introduction into the discourse of a referent by means of a pronoun or a zero rather than the expected NP.

Apart from language specific points as discussed above, Givon's account of topic continuity is generally consistent with patterns of reference discussed later. He summarises his approach in Gricean terms which reflect Brown and Yule's observation about the lexical attenuation of given entities:

'The more disruptive, surprising, discontinuous or hard to process a topic is, the more *coding material* must be assigned to it.'

and:

'Expend only as much energy on a task as is required for its performance' (Givon 1983: 18).

Chafe (1980), presents a strong argument for a non-linear, hierarchical description of anaphoric reference which relates to human consciousness. He pinpoints the centre of interest in a text or dialogue and with it the focus of consciousness. He suggests that speech is produced in spurts, called idea units, and that a series of idea units may be a centre of interest (Chafe 1980: 48). Humans use a focus of consciousness, based on learned schemata, intellect and judgement, to scan the centre of interest. This theory, more fully discussed above in Section 1.3 has given rise to theories of anaphoric reference based on hierarchical structures, such as those of Chen (1986), Fox (1987), and Tao (1996) as described below in Sections 4.3.1.2, 4.3.2.2.3, and 4.3.2.2.8.

#### **4.2.2.1 Topic-comment in Chinese in both spoken and written texts**

Many Chinese linguists, while differing slightly in their interpretation, consider topic-comment structure to be the basic structure of spoken and written Chinese, rather than subject-predicate structure. Topic and comment are related semantically or pragmatically:

26.

|   |        |     |         |
|---|--------|-----|---------|
| 北京                                      | 故宫     | 最   | 出名      |
| Beijing                                 | gugong | zui | chuming |
| Beijing Imperial Palace most well known |        |     |         |

(as for the palaces in ) Beijing, the Imperial Palace is the best known.

Li and Thompson, in their study of anaphora in written texts, claim that 'there are no structural properties predicting the interpretation of the referent for zero-pronouns but that the interpretation of the referent for unrealised pronoun is inferred on the basis of pragmatic knowledge' (Li and Thompson 1979).

Tao Hongyin, among others, defines the topic in terms of 'aboutness' (Tao 1996: 106). He subcategorises sub-types of NP which 'relate to the various aspects of the dynamic process of referring: referent introducing, activating, framing, anchoring, reinforcing, predicting, listing, topicalization and contrasting' (Tao 1996: 100). La Polla points out that the topic in Chinese is associated with givenness, identifiability, definiteness and accessibility (la Polla 1995).

Huang, in his study of zero anaphora in conversation, notes that while the topic comment construction forms the basis of Chinese grammar it may be ambiguous, for, as noted above, the listener or reader may be dependent on context, shared knowledge or background knowledge to interpret it (Huang Yan 1994: 163). This reflects Givon's viewpoint that a less than perfect correlation exists between code and message, and it is context that allows syntax to be an efficient processing device (Givon 1983: 16).

The roles of subject and object are catered for within the topic-comment framework. Tao Liang (1996), investigating spoken discourse, provides the following definition:

*'A subject is the NP..... that has a grammatical relationship with the predicate verb in the following fashion: it is one of the main arguments, usually the agent , of a transitive verb in the active voice , or the only argument of an intransitive verb (Chao 1968; Ding et al. 1979). An object is one of the main arguments of a transitive verb that is the receiver of the action of the verb in the active voice. ....A topic..... refers to an NP referent that is the center of a discussion in the discourse' (Tao 1996: 488).*

#### **4.2.2.2 Topic-comment as syntax**

Shi (1989) has argued strongly that the topic-comment structure in Chinese is not discoursal, but syntactic. He claims that the topic has all the syntactic functions normally assumed for an S in Chinese, as it may be the subject, the verbal complement, the NP complement, the adverbial and the modifier of an NP within a Chinese sentence and having its own syntactic properties, is a separate category in the grammar (Shi 1989: 223). He notes that the domain of a topic may stretch across several sentences, each of which function as a comment of the topic, have some kind of 'aboutness' relationship, and contribute new information. Crucially, the initial NP topic together with any one of the comment sentences constitute an independent topic comment. 'Such a topic can license deletion of the topic of all the other sentences under the rule of identical topic deletion' (Shi 1989: 224). This description differs from that of Li and Thompson (1979), whose explanation is that the null arguments in the subsequent sentences are zero pronouns referring anaphorically to the initial topic and that resolution depends on pragmatic features. According to Shi a zero anaphor in the subsequent sentences is permissible as long as the initial NP is in subject or object position, or is a specifier of the subject NP (Shi 1989: 226). Shi's example, below, is very typical of Chinese sentences in which object (new) in the first part of the sentence becomes subject (given) in the second or subsequent parts. In this example,

the topic is object NP of the first sentence and the second sentence, then becomes the subject (of a passive verb) in the third sentence:

27.

我看上了这个姑娘，他也看上了，最后被他抢走了

wo kanshang le zheige guniang, ta ye kanshangle, zui hou bei ta  
qiangzoule.

I fancy (asp.) this (measure) girl, he also fancy (asp.), finally (zero  
pronoun) (passive) he steal (asp.)

I fancied this girl, he fancied her too, in the end he stole her.

Note that the 'sentences' in this example are not separated by full stops and might be regarded by some as clauses, as described above in Section 2.5.7.4.

Shi also addresses the question of definiteness, maintaining that the first mention NP is usually indefinite and becomes definite in subsequent mentions (Shi 1989: 228). He also notes that some predicates can take a whole topic chain as a subject, a phenomenon rather similar to English, as in the following example:

28.

李四不喜欢念书，总是看电视，成绩越来越差，使他妈妈最伤心的事

Lisi bu xihuan nian shu, zong shi kan dian shi, chengji yue lai yue cha,  
shi ta mama zui shangxin de shi.

Lisi not like study, always watch television, results increasingly lack,  
cause his mum most distress (mod.) thing



Lisi doesn't like studying, (and) always watches TV, (and) his grades get worse and worse, (a fact) which distresses his mum.

Another reason for regarding topic-comment as a syntactic rather than discursal structure, according to Shi, is that the topic chain can function as the complement of a NP. He goes on to question whether topic chain is not just another term for co-ordinate construction, especially since conjunctions are often omitted in Chinese, a point well illustrated by the example above. Shi suggests that the topic chain should be treated as the largest unit in Chinese syntax, S, while single link topic chains should be regarded as a subset of S (Shi 1989: 254).

#### **4.2.2.3 The problem of zero topic**

C.-T. J. Huang (1989), among others, posits the notion of empty topic for Chinese, in other words, the topic may be realised initially as zero on the basis that inflection in Chinese is lexical, and therefore inflection in Chinese can still govern the topic and the subject. Huang Yan disputes this by saying that the empty topic is a discourse rather than a sentence phenomenon and a topic can only be empty if the chain initial topic is present (Huang Yan 1994: 46). Somewhere in the discourse, the topic will have been referred to by means of a full NP, but the common phenomenon of return pop, discussed later in Section 4.3, may give the appearance of a null topic. In addition in Chinese discourse there may be a null argument which has no antecedent. This type of null argument may be similar to an expletive pronoun in other languages (eg 下雨 XIA YU it's raining), or may be deictic - the topic is in the non-linguistic context - or may be equivalent to an English 'dummy subject' reflected in the very common YOU 有 (there is/are) in Chinese, see Footnote 14.

### 4.3 Factors affecting speakers' or writers' choice of full NP, pronoun or zero pronoun

#### 4.3.1 Anaphoric choice in languages other than Chinese

The term anaphoric choice is used to refer to speakers' or writers' selection of anaphors, and how or why these selections are made.

Analyses of the factors determining the use of zero pronouns as anaphoric devices differ according to the ideological background of the researcher. However, as noted previously in Chapters 2 and 3 on reading and translation, certain typological characteristics of Chinese point towards a more discourse based interpretation of the use of null arguments in all grammatical environments, in all types of text and in both written and spoken modes.

##### 4.3.1.1 Government and binding theory approaches to anaphora in languages other than Chinese

Chomsky's distinction of Anaphors and Pronominals lies in the requirement that an anaphor must be referentially dependent and bound in a syntactic domain; a pronominal, however, may be referentially dependent but free in a syntactic domain. This gives rise to a typology of four types of nominal expression including both overt and non-overt, as follows:

Chomsky's typology of NPs

|                               | lexical         | empty      |
|-------------------------------|-----------------|------------|
| a. [+ Anaphor, - Pronominal]: | lexical Anaphor | NP-trace   |
| b. [- Anaphor, + Pronominal]: | pronoun         | <i>pro</i> |
| c. [+ Anaphor, + Pronominal]: | -               | PRO        |
| d. [- Anaphor, - Pronominal]: | name            | variable   |

The column on the far right of the table above encapsulates Chomsky's typology of empty categories (EC).

Conditions for the category of referential EC include the Projection Principle, according to which lexical properties should be projected on to all levels of syntactic representation, and the Extended Projection Principle, according to which every sentence must have a structural subject. NP-trace and variable (above) are subject to the Empty Category Principle which states that traces must be properly governed.

#### **4.3.1.2 Discourse based approaches to anaphora in languages other than Chinese**

Fox (1987) in her discussion of anaphora in English, takes the view that while a text, spoken or written, may be linear, the design and comprehension of a text is hierarchical. Her framework, based on a pattern of discourse units, contrasts with traditional theories of anaphora such as that of Givon, above, which hold that the anaphoric devices selected by the speaker or writer are related in terms of proximity to antecedents (Fox 1987: 18), but is consistent with the views of Chafe (1980). Fox believes that the linear theory does not account for a critical portion of her data though her findings seem to suggest that while distance from the antecedent does not affect pronominalisation in conversational discourse, there seem to be some limitations on pronominalisation in written expository prose which are imposed by distance from the antecedent. For both types of discourse, however, other patterns apply.

Fox suggests that while a sequence is opened with a full NP, the subsequent use of pronoun or pronouns shows that the speaker understands that the sequence has not closed down. The type of anaphoric device chosen reflects monitoring for a sign of closure.

#### 4.3.1.2.1 Rhetorical structure analysis as applied to anaphora

In her analysis of anaphora in English expository texts Fox uses rhetorical structure analysis as a model for describing a range of argument relations and at the same time showing texturing within a text, ie, capable of capturing the relative centrality or peripherality of discourse units in a text. Fox maintains that texts are

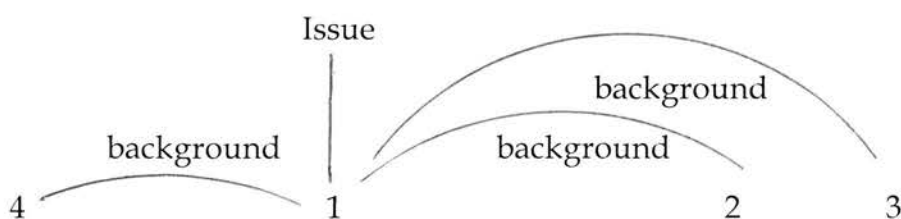
*'groups of hierarchically organized clauses which bear various informational and interactional relations to one another.....the model (rhetorical structure analysis) has in its apparatus a basic unit - the proposition - and a class of text structures which describe the structures which the propositions display' (Fox 1987: 78).*

In Fox's framework the smallest unit of text is the proposition. Propositions are grouped into a unit called the R-structure. An R-structure in most cases consists of a nucleus and an adjunct, but in some cases, such as a text which is a list, may consist of a series of nuclei, or a nucleus and one or more adjuncts. The nucleus may, for example, be an Issue, and adjuncts may provide evidence or elaboration or background information, for example:

29.

- (1) John Smith has been appointed Head Teacher of St. Boisel's School.  
 (2) He has taught in the school for fifteen years and (3) has been acting head since the retirement of Miss Jones. (4) He is an active member of the NUT.

Fox illustrates this type of Issue in diagram form:



(Fox 1987: 80)

Fox points out that any part of an R-structure can be realised by another embedded R-structure.

#### **4.3.1.2.2 Rhetorical structure analysis as applied to anaphora in English**

Fox bases her argument on the following assumption:

- '1. Anaphoric form X is the unmarked form for a context like the one the participant is in now.*
- 2. By using anaphoric form X, then, the participant displays an understanding that the context is of that sort, then the other parties may change their understanding about the nature of the context to be in accord with the understanding displayed' (Fox 1987: 93).*

Fox assumes that these conditions, formulated to deal with spoken discourse, are equally appropriate to written discourse, in that the writer anticipates the reader's comprehension and the reader attempts to understand the writer's intention, a view which has a bearing on any theory of comprehension of null arguments or zero pronouns.

In this sense, for the producer of the text, whether spoken or written, context may determine use and use may determine context: by choosing to use a pronoun, rather than a full NP, the speaker/writer shows the listener/reader that the referent is in an 'active' or 'controlling' position. The adjunct of an R-structure is active while its nucleus is being produced and the nucleus of an R-structure is active while its adjunct is being produced. A proposition is controlling while its R-structure partner is active (Fox 1987: 93-96).

#### **4.3.1.2.3 Return pop in English**

Fox notes that a physically distant R-structure can 'tie' back to an earlier proposition, a proposition other than the immediately preceding one. She terms this phenomenon 'return-pop' after Reichman 1981 (Fox 1987: 18):

30.

But as he was approaching 60 a chance encounter with a humanist psychologist, Dr. Carl Rogers, made *him* think about a change. (Los Angeles Times, July 11, 1983, cited in Fox 1987: 98)

#### **4.3.1.2.4 Possible differences in anaphoric choice in conversational and written genres**

Fox's findings appear to suggest that while in conversation a topic may not be closed for a considerable time, and therefore 'return pops' - the continued use of a pronoun rather than a full NP to refer to a reintroduced referent - may be permitted over a number of discourse units, in English expository prose a return pop may only 'pop over' material which contains mention of the relevant referent. While return pop pronouns may be used over deeply embedded material in conversation, a return pop in expository prose is only permitted over one level of embedding. Written text allows physical traces of closed off material to remain, for the reader to trace at leisure, whereas closure in conversation is more final, and thus the topic may need to remain active for longer (Fox 1987: 140 -142). It is important to note that different genres, and discourse in different languages could vary in the pattern of pronominalisation.

Fox's model of description could apply appropriately to patterns of zero anaphoric choice in Chinese, as noted by Li and Thompson, see Section 4.3.2.2. below. In Fox's descriptive framework of English, the normal pattern is a full NP followed by a series of pronouns, while the comparable pattern in Chinese would be a full NP followed by a series of zeros.

#### **4.3.1.2.5 A discourse based view of anaphora in Japanese**

A study by Clancy (1980) focusing on anaphoric choice, comparing Japanese and English using the pear stories, is perhaps one of the most relevant for

Chinese. The pear stories are a corpus of spoken retellings in various languages of the plot of a film about the theft of some pears (Chafe 1980). Like Chinese, Japanese favours zero for anaphoric reference but tends to use NP rather than pronoun for explicit anaphoric reference.

Taking into account the speaker's awareness of the addressee, Clancy notes that cognitive constraints on anaphoric choice may include factors of time and interference, that is, the time since the previous mention (comparable to distance or recency in a text) and the number of other referents mentioned in between (Clancy 1980: 129). While finding that NP reference differed a great deal between English and Japanese, Clancy found that distributions of English pronouns and Japanese ellipsis (zero) were quite similar with regard to time and interference. She concluded that 'distributions of inexplicit reference forms are basically the same in both languages' (Clancy 1980: 140). Speakers of both English and Japanese were found to use inexplicit reference for characters who had already been introduced in the narrative. English speakers were likely to use a pronoun for reference across three or fewer sentences, Japanese speakers were likely to use zero for reference across three or fewer clauses. Except in a very few cases in each language, a NP was used to introduce a new referent or was used anaphorically at an 'episode' boundary. Clancy notes that the shift of reference type at an episode boundary could either be a listener oriented strategy, or a speaker based strategy having to do with memory constraints, a factor which will be further discussed in Section 5.9.

However, she found that Japanese speakers tended to use NPs for anaphoric reference where it was not cognitively necessary. This may be explained by the notion of empathy (Kuno and Kaburaki 1977): 'In describing a particular event or state...a speaker may make certain lexical and syntactic choices which indicate that he is identifying with one of the characters being referred to'



(Clancy 1980: 179). In such cases, the speaker generally uses a pronoun or zero for the person being empathised with, and a full NP for characters who do not attract the speaker's empathy. Kuno and Kaburaki postulate an 'empathy hierarchy' which may help to explain unexpected choices of anaphoric reference. Their 'surface structure empathy hierarchy' indicates that one is most likely to empathise with the referent of the subject, then with that of the object and so on (Kuno and Kaburaki 1977). Waggot (1997a) found that in dialogue in Chinese, terms of address became more prolix not only in situations of greater formality, but also in situations of great familiarity combined with great hostility, a phenomenon which may point to more explicit expression for intensity of emotion.

Clancy concludes that individual variation and universal cognitive constraints play a part in anaphoric choice. Her findings echo those of Li and Thompson (1979), supporting tendencies in anaphoric choice rather than hard and fast rules.

#### **4.3.1.2.6 A compromise view of anaphora in English**

Reinhart (1983), refuting her own previous (1976) analysis, but basing her revised conclusion on the same empirical evidence, argues that a distinction should be made between pragmatic coreference and the interpretation of pronouns as bound variables, restricted by sentence level rules (Reinhart 1986: preface). She argues that while the linear order of constituents plays no part in grammatical theory, it is crucial in pragmatic theory, that to refer back to a newly named referent is the natural tendency of coherent discourse (Reinhart 1986: 37). She claims that it is impossible to tackle coreference in purely discourse terms, but that there is a significant correlation between syntactic relations and semantic properties.

However, Reinhart's study is based on intra-sentential anaphora alone and cannot deal with ambiguities and misunderstandings which arise in longer pieces of discourse.

### 4.3.2 Anaphoric choice in Chinese

#### 4.3.2.1 Government and binding theory approaches to anaphora in Chinese

C.-T. J. Huang has suggested modifications to Chomsky's model (outlined above in Section 4.3.1.1) which he feels to be more appropriate to Chinese. He proposes that PRO and *pro* should be brought together into a single Pronominal Empty Category, which he calls *pro*/PRO. His typology can be tabulated as follows:

- a. [+Anaphor, - Pronominal]: NP-trace
- b. [-Anaphor, + Pronominal]: *pro*/PRO
- c. [-Anaphor, + Pronominal]: variable

C.-T. J. Huang posits a Generalised Control Rule according to which an empty pronominal is controlled in its control domain if it has one. He further suggests an empty topic analysis, a somewhat problematic area, since, as Huang Yan demonstrates, infinite regression would be required to cope with an empty topic, as discussed in Section 4.2.2.1. C.-T. J. Huang's analysis, in common with other GB approaches, depends on proximity, requiring a pronoun to be identified by its closest subject (C.-T. J. Huang 1982) and later proposing a Generalised Control Rule which states that an empty Pronominal is coindexed with the closest nominal element (C.-T. J. Huang, 1984). Huang Yan demonstrates, by means of a very common structure in Chinese, that this analysis is inappropriate:

31a.

病人 说 医生 知道 明天 给他 开刀

bingren shuo yisheng zhidao 0 mingtian gei ta kaidao.

patient say doctor know tomorrow for 3SG operate.

'The patient says the surgeon knows (I/you/he/she/we/they) will operate on him tomorrow.'

31b.

医生 说 病人 知道 明天 给他 开刀

yisheng shuo bingren zhidao 0 mingtian gei ta kaidao.

doctor say patient know tomorrow for 3SG operate.

'The doctor says the patient knows (I/you/he/she/we/they) will operate on him tomorrow.'

As Huang Yan points out, a GB type approach depending on proximity would result in a counter-intuitive interpretation of (b). The only sensible interpretation of (b) is based on our knowledge of the world - that surgeons operate on patients and not vice versa, so the likely interpretation of the null argument is he or she, ie. the doctor, or they, ie. the doctors.

#### 4.3.2.2. Discourse based approaches to anaphora in Chinese

Li and Thompson (1981) in their reader friendly grammar of Chinese point out that nothing is actually missing from a Mandarin sentence which has a null argument:

*'One common situation in which noun phrases are unspecified is the topic chain, where a referent is referred to in the first clause, and then there follow several more clauses talking about the same referent but not overtly mentioning that referent' (Li and Thompson 1981: 658).*

This is a very clear description of the most straight forward pattern of null arguments in Chinese. They go on to say that the zero pronoun does not

necessarily have to refer to the topic of the preceding clause, but perhaps to the secondary topic, subject or object.

They claim that in their examples the 'sense of the passage makes it clear' to what or whom the null argument refers. Their criteria for choice of pronoun rather than zero are firstly the guideline of highlighting and secondly information that is unexpected or unusual (Li and Thompson 1981: 662).

Earlier work by Li and Thompson (1979) gives a more detailed explanation of the distinction between pronoun and zero pronoun in Chinese. They based their study on texts taken from two novels written in modern Chinese, but not written in the twentieth century. Taking 50 Chinese reading informants, from Taiwan and California and in the case of both novels including the author as one of the subjects, they asked their subjects to insert overt pronouns or zero pronouns in a number of texts at appropriate points. They found that there was a wide range of disagreement as to the correct choice of zero or overt pronouns. They concluded that

*'zero-anaphora.....in Chinese.....is so widespread and so unconstrained that it must be considered as the normal mode of NP-anaphora in Chinese, the interpretation of referents for zero pronouns being inferred on the basis of semantic and pragmatic knowledge and of information present in the discourse itself' (Li and Thompson 1979: 334).*

Where the use of the overt pronoun was concerned they came to the conclusion that connectedness or 'conjoinability' appeared to be the factor most likely to affect selection of zero pronoun rather than pronoun. Their criteria for identifying 'conjoinability' seem rather vague and subjective, though the criteria for non-conjoinability seem more plausible; for example they suggest that contrastive clauses are non-conjoinable and therefore require the selection of an overt pronoun. This is in line with their later highlighting or unexpectedness guidelines. In summary, what emerges from

Li and Thompson's early study is the importance of pragmatic factors in choice of anaphoric reference; structural properties are unlikely to affect the choice of pronoun type and choice will vary from speaker to speaker and within speakers depending on perception of the pragmatic situation (Li and Thompson 1979: 312).

#### **4.3.2.2.1 Relative frequency of nouns, pronouns and zero pronouns in Chinese**

Li and Thompson's notion of the anaphoric zero pronoun as 'the normal mode of NP- anaphora in Chinese' (Li and Thompson 1979) is supported by Xu (1990), who made a study of the distribution of the 3rd person singular pronoun TA 他 / 她 ( *he* or *she* ). Inevitably this study also covered the distribution and relative frequency of nouns, pronouns and zero pronouns.

Xu states that Miao (1987) and Chen (1987) both claimed that in subject position in a sentence the zero pronoun was more frequent than third person singular pronoun and shows that his data yield the same result. Following Givon (1983) and Li and Thompson (1979), he maintains that the high frequency zero pronoun is a measure of topic continuity. The pronoun is secondary in maintaining topic continuity, while the NP is weakest in maintaining topic continuity (Xu 1990: 328). Within a discourse analysis framework, however, Xu points to the factor of proximity, noting that because the zero pronoun has no overt form, the addressee must make decisions on its referent and the speaker/writer must ensure that sufficient information is provided; this requires that the zero pronoun be located close to noun or pronoun, and that long and complex constituents should not be placed between the zero pronoun and its referents (Xu 1990: 328).

The relative strength and weakness of the noun and pronoun in topic continuity reflect categorisation in the human brain, according to Xu, setting

up a relatively abstract concept which can only create an overt, 'real' image if connected to a noun (Xu 1990: 328). This notion may bear some relation to the theory of mental spaces discussed in Section 1.4 and will be taken up again in relation to my data in Chapters Five and Six.

Xu points out that while in a text with one protagonist, first mention tends to be a noun and subsequent mentions tend to be pronominal, in a text containing more than one protagonist, not only first mentions, but also subsequent mentions may be nouns, in order to avoid difficulty for the reader. He suggests that not just new protagonists, but also new episodes require nouns (as Clancy suggests, see Section 4.3.1.4 above), rather than pronouns or zero pronouns (Xu 1990: 333). Xu also suggests that time adverbials tend to be followed by a pronoun, but his data are inconclusive.

Xu's frequency data show that overall, where there is a pause between clauses (manifest as a comma or full stop in writing) the anaphor in the second clause is most likely to be a pronoun, or secondly, a noun. However, in situations where there is no pause between clauses, the anaphor is likely to be a zero pronoun, again showing the stronger role of the zero pronoun in topic continuity. The latter notion may reflect Li and Thompson's notion of conjoinability (Section 4.3.2.2 above), though they do not mention punctuation as a criterion.

#### **4. 3. 2. 2. 2 A centering based refutation of Li and Thompson's discourse analysis view**

In an experiment based on centering theory, Yang, Gordon, Hendrick and Wu (1999) specifically attempted to refute the discourse based theories of Li and Thompson (1979), Givon (1983) and Chen (1986) which argue that anaphoric choice and resolution are determined by pragmatic and contextual factors. Yang et al. state that 'the context-and-pragmatics-only approach advocated by

Li and Thompson..... is not consistent with characteristics of anaphoric reference in other languages, where referential processing is found to be influenced by syntactic factors' (Yang et al. 1999: 719).

Their research replicated experiments on English which had used response times, coherence and grammaticality judgements, reading time and eye tracking. They state that 'studies of English have shown that reduced referential expressions .... contribute more to discourse coherence than do unreduced expressions' (Yang et al. 1999; 715). In their experiments they found that processing of coherence in Chinese discourse was similar to that of English, and that overt pronouns and zero pronouns contributed equally to coherence in Chinese discourse, a finding in marked contrast to that of Xu, above (Yang et al. 1999: 715).

Their subjects (all psychology undergraduates and native speaker/readers of Chinese) were timed in reading the stimuli electronically and were required to answer a true/false comprehension question. They were given feedback only if their comprehension answers were incorrect.

Noting that Gordon et al. (1993) had found a repeated name 'penalty' in English where reading time for sentences with repeated names was greater than for sentences with pronouns, they predicted that readers of Chinese would also pay a 'repeated-name penalty.' Secondly, they tested the distinct roles in coherence of overt pronouns and zero pronouns. Their position was that if the traditional view of the Chinese zero pronoun as making the strongest contribution to discourse coherence was correct, the zero pronoun should make discourse easier to understand (Yang et al. 1999: 721).

They found that, as in the English experiments, reading times were slower for sentences containing repeated name referring expressions in subject position,



but there was an insignificant difference in reading time when the repeated-name referring expression was in object position. Similarly, there was a greater, but not significantly so, reading time required for sentences containing overt pronouns as opposed to zero pronouns. They take pains to note that longer reading times are not, as some of their critics suggest, due to additional material.

Yang et al. conclude from their results that reduced referring expressions (overt and zero pronouns) can facilitate discourse comprehension, as in English. Furthermore, they conclude that the insignificant difference in reading time for overt pronouns and zero pronouns indicates that 'there is no need to assign special status to the zero pronoun in Chinese' (Yang et al. 1999: 725).

My view is that Yang et al.'s experiments in fact support the context and pragmatics theories of Givon, Li and Thompson and Chen. The major difference in reading time required for the repeated name in subject position condition, the slight difference in reading time required for the repeated name in object position condition and overt pronoun condition reveal a pattern of topic continuity as outlined by both Givon and by Li and Thompson.

Apart from this, there are, to my mind, some anomalies associated with Yang et al.'s experiments. Firstly, they seem to have mis-apprehended the thrust of Li and Thompson's argument. Zero pronouns are not, as Yang et al. state 'special' or 'unique' and do not make a 'stronger' contribution to discourse; rather, they play a role which is distinct from those of nouns and overt pronouns - each type of referring expression contributes equally, but differently, to discourse.

Secondly, they tested short texts of a uniform pattern which involved proper names - they did not test common nouns; they tested only animate subjects and objects; and they took care to avoid confusion of referents of the same gender. In the comprehension section of the test they offered a forced choice of two options - a method which is now widely disused owing to its unreliability. Thus, they have not only disguised a pragmatic-contextual notion in a grammatical cloak, but they have also neatly avoided all the problems usually associated with anaphoric resolution in Chinese.

#### **4.3.2.2.3 Topicality hierarchy in reference tracking**

P. Chen (1986) in his discourse analysis study of referent introduction and tracking in Chinese notes that Chinese, like for example, Japanese and Thai, is a language whose reference maintenance systems rely on inference and which exhibits abundant zero anaphora. While Chen notes that referent introducing and tracking vary as to individual speakers and their goals, he claims that anaphoric choice in this type of language is not as random as it may seem.

Chen used the Pear Stories (Chafe 1980) to collect his data. He found that about 70% of zero anaphors (ZA) occurred in the same sentence as their referents and about 50% of pronominal anaphors (PA) occurred across a single sentence boundary, finding a correlation (though not a strong correlation) between zero anaphors and same sentence reference, and pronominal anaphors and sentence boundaries. He suggested also that another factor in anaphoric choice might be interfering referents. Interfering referents occur between the anaphor and its antecedent. They are manifest as NPs or pronouns and could be interpreted as antecedents of the zero anaphor. In cases where there are interfering referents there is less likelihood of a zero anaphor. This, in Chen's view, supports Clancy's notion that there are cognitive constraints on the range of referential forms an addressee can

process. These constraints may include time passed since last mention of a referent and the number of other interfering referents (Clancy 1980: 129).

Chen claims that a further factor in anaphoric choice is persistence - that is the continued presence of a semantic argument over a number of subsequent clauses. Chen found that zero anaphors were shorter in persistence than a pronominal anaphor. He notes that while distance and interference may be acting as parameters which measure this continuity, however, distance and interference are derived from other more fundamental factors of discourse hierarchy and are not in themselves the most basic factors affecting continuity (Chen 1986: 127).

Chen proposes a topicality hierarchy for Chinese:

PREVBL 1 > POSTVBL 1 > PIVOTAL OBJ > OBLIQUE OBJ > POSTVBL 2

(preverbal) (postverbal)

[ ZA (zero anaphor).....]

[ PA (pronoun anaphor).....]

[ NA (noun anaphor).....>>] (Chen 1986)

He concludes that the more attenuated forms tend to occur in syntactic slots high in the hierarchy. Thus, the most topical, or central, and persistent argument is more likely to occupy the preverbal position and be manifest as a zero anaphor. A less topical, central, less persistent argument is more likely to occupy postverbal or pivotal positions and be manifest as a pronoun or noun anaphor.

Chen’s proposed hierarchy may be seen to complement Givon’s (1983) hierarchy (see Section 4.2.2 ) and Comrie’s (1981) hierarchy:

AGT>DAT>GEN>PAT>LOC

As animacy is seen to be one of the properties related to genitivity, this semantic case hierarchy may be seen to complement the animacy hierarchy:

HUMAN>ANIMATE>INANIMATE>ABSTRACT>

But Chen reminds us that 'Chinese displays a marked reluctance to encode inanimate reference by means of pronominal anaphora' (Chen 1986: 137).

He points out cases in his data which appear contrary to the hierarchies he has outlined; attenuated (PA or ZA) anaphors where continuity is not high, or conversely, NA (noun anaphors) where there are no competing referents and no risk of ambiguity. He rationalises this by citing Givon: 'the more disruptive, surprising, discontinuous or hard to process a topic is, the more coding material must be assigned to it' (Givon 1983). Chen found that addressees tended to interpret zero anaphors as anaphors of the prior subject in the preceding clause. This he regarded as a default solution, when not semantically incompatible. Explicit anaphors (pronouns or nouns) are to be expected when the reference of the subject of the prior clause is discontinued. However, Chen found that zero anaphors were used to encode when reference switched from a prior subject to a new one, but that this phenomenon occurred more in writing than in speaking. This is an observation developed by Tao (reviewed below in Section 4.3.2.2.7), who found that in Chinese, new, noteworthy information may be introduced, or discontinued information reintroduced by means of a zero anaphor. This also reflects the observations by Clancy (above, Section 4.3.1.3) and Xu (Section 4.3.2.2.1). This type of zero anaphor is well known for its ambiguity and since it carries no information the addressee must use some other means to identify the referent. Chen notes that the selectional restrictions of the predicate may give essential clues as to the referent. He gives the example of CHU JIA 出嫁 (to be married (off)) which can only apply to a female, but this could still be ambiguous as to referent, if more than one protagonist in the text is female. The importance of the verb in determining the referent of a zero anaphor is

adumbrated by Chen and more fully discussed by Tao, and is more fully explored in relation to the data in Chapter Five.

Chen does raise the issue of parallelism in Chinese, noting that similar syntactic configurations, often containing lexical, semantic or rhetorical parallels may provide clues to anaphoric resolution, as the first of a pair of clauses may provide background information for the second. Zhao (2000) found that in parallel sentences used in advertisements (but of a type very common in Chinese writing) there was a distinct theme rheme structure, in which the first rheme became the second theme and so on, in a similar pattern to that seen in sentences with overlapping reference control, where the object of a previous clause becomes the null argument subject of the next clause. Thus, the topic comment structure of Chinese, whether regarded as syntactic (Shi 1990) or discoursal (Li and Thompson 1981) provides a pattern of closely linked given and new information.

Chen summarises by saying that anaphoric choice and resolution may be subject both to linear and hierarchical perspectives. The application of distance and interference theories should be sensitive to hierarchical structure, but 'hierarchical closeness at the same hierarchical level overrides the linear distance and intervention' (Chen 1986: 180).

#### **4.3.2.2.4 Negligibility**

Chen states that not all the instances of ZA in his data can be explained in terms of high continuity of referents. He posits a theory of high negligibility of the referent as yet another factor in anaphoric choice. For example, referents at paragraph initial boundaries may be considered low in continuity, but may be realised as ZA. Moreover the animacy hierarchy, for instance, indicates that inanimates are less noteworthy than animates, and adjunct clauses are less important than nuclear clauses, a notion in line with principles

of economy and efficiency and appropriate to considerations of short term memory span (see Section 2.2.).

#### **4.3.2.2.5 Pronouns in a topicality hierarchy in reference tracking**

Having discussed ZA at length, Chen devotes some space to a theory of the occurrence of PA and NA. In his data the 126 pronoun subject referents are all found to be coreferential with the subject in the prior clause, but separated from the preceding clause by sentence boundaries. According to Chen 88% of these cases showed a change of schema and a different goal across the sentence boundary, as compared with cases where a single schema or goal is maintained and ZA is used. He claims that the choice of PA can be explained in terms of high noteworthiness of the referent.

But he concedes that PA may be used for a switched reference subject, or could be in free variation with ZA: 'In the case where the PA by itself is not able to distinguish the anaphor from competing candidates in the context, it may still be chosen when other linguistic or extra linguistic information leads to easy identification of the antecedent' (Chen 1986: 216). He notes however, that ZA may also be used when the form of PA is not sufficiently differentiating. Chen summarises the factors for the choice of a pronoun rather than a zero as follows:

- '1. It occurs in places which are marked by minor breaks in discourse continuity and
2. It stands high in saliency in discourse' (Chen 1986: 222).

#### **4.3.2.2.6 Noun anaphora in topicality hierarchy in reference tracking**

Chen's view is that the greater the distance between referent and anaphor the more likely the choice of noun anaphora. But he admits that there are 'quite a few' instances which cannot be explained in this way and suggests that a

break of coherence in, chiefly, topic, and also in event, time and space, may account for the choice of a noun.

#### **4.3.2.2.7 A Gricean account of zero and pronominal anaphora in Chinese**

Huang Yan (1994) takes a Gricean view of anaphora, arguing strongly that the extent to which syntax and pragmatics interact varies according to language typology, and that in some languages, such as Chinese, Japanese, Korean, etc., pragmatics may play a more central role than grammar, (Huang 1994: 2) a point also made by Chen above.

Huang points out that Grice's conversational implicatures are non-logical inferences that 'contain messages which are meant without being said in the strict sense' (Huang 1994: 4). This is of particular relevance to Chinese discourse, considering the predominance of ellipsis and implicitness found in Chinese, discussed in Sections 2.5.7, 3.5.1 and throughout this chapter. Huang bases his theory of anaphora in Chinese on Levinson's Q- I- and M-principles, a later development of the Gricean principles of Quality, Quantity, Relation and Manner. A crucial aspect of Levinson's framework is the addition to the 'speaker's maxims' of the 'recipient's corollary.'

Levinson's principles are as follows:

##### **1. The Q-principle ( quantity):**

The speaker should not provide a statement informationally weaker than his/her knowledge of the world allows, unless providing a stronger statement would contravene the I-principle.

The recipient should assume that the speaker has made the strongest statement consistent with what he knows.

##### **The I-principle (information):**

The speaker should say as little as necessary to achieve communicational ends.



The recipient should amplify the informational content of the speaker's utterance by finding the most specific interpretation. This includes the notions of assuming stereotypical relations, assuming the existence or actuality of what a sentence is about, and to assume referential parsimony - that is, to prefer coreferential readings of pronouns or zeros.

The M-principle (manner):

The speaker should not use a prolix, obscure or marked expression without reason.

The recipient should assume that if the speaker uses a prolix or marked expression he is trying to avoid stereotypical associations (I-implicatures) (Levinson 1987).

Huang states that the Q- and I- principles operate in opposite directions: while Q-implicatures decree that if an informationally stronger expression is absent, one infers that the interpretation associated with the use of that expression does not hold, I-implicatures indicate that a semantically general expression requires a semantically specific interpretation (Huang 1994: 9-10). Huang's examples show clear evidence that this type of inference is more informative than the utterance which generates the inference. This is obviously of great relevance to Chinese where zero anaphora is considered to be the normal mode of NP-anaphora (Li and Thompson 1979). According to Huang the inferences induced by the I-principle are optimally in line with our knowledge of the world.

Applying the Q- I- and M- principles to anaphora in Chinese, Huang points out that 'the use of a reduced, semantically general anaphoric expression tends to favour a local coreferential interpretation, whereas the use of a full

semantically specific anaphoric expression tends to favour a local, non-coreferential interpretation.

In order to account for anaphora in Chinese, assuming the semantic content hierarchy, Huang adopts the application of the I- and M- principles and adds consistency constraints of disjoint reference, presumption, information saliency and general implicature constraints as follows:

*'the use of a zero anaphor will I-implicate a local co-referential interpretation; the use of a pronoun will I-implicate a local co-referential interpretation, unless the pronoun is used where a zero anaphor could occur, in which case, the use of the pronoun will M-implicate the complement of the I-implicature associated with the use of the zero anaphor; ...the use of a name or a lexical NP where a pronoun or a zero anaphor could occur , will M-implicate the complement of the I-implicature associated with the use of the pronoun or the zero anaphor, in terms of either reference or expectedness'* (Huang 1994: 17).

These principles reflect Li and Thompson's views that zero pronouns are the normal mode of anaphora and that a pronoun or a NP would be used anaphorically when there is an element of unexpectedness or a need for highlighting - possibly manifest in 'non-conjoinable' contrastive clauses as discussed in Section 4.2.2.1.

Huang imposes consistency constraints as follows: an interpretation resulting from the framework outlined above must be consistent with the Disjoint Reference Presumption, that is 'the arguments of a predicate are intended to be disjoint , unless marked otherwise' (Huang 1994: 17). ( The DRP is strongly reminiscent of Chomsky's comments on disjoint reference binding condition, but Huang insists that the DRP is of a pragmatic nature. It describes usage preference especially in the case of Chinese.) Interpretations must also (a) be consistent with information saliency, that is, higher constructions take precedence over lower constructions, and (b) in line with a hierarchy of topic>subject> object>etc. (Huang 1994: 17).

Huang argues that 'a proposition A is more informative than a proposition B if and only if the set of entailments of B is properly contained in the set of entailments A' (Huang 1994: 121). To illustrate his point that the fewer existential commitments there are, the more informative the proposition, he cites the following example in English:

32.

John adores his wife.

In this case a coreferential interpretation is more probable.

Huang claims that his neo-Gricean theory of anaphora will describe many of the patterns of intrasentential anaphora in Chinese and can be extended to account for anaphora in Chinese conversation, provided that principles of politeness and organisation of conversation are taken into account (Huang 1994: 18).

#### **4.3.2.2.8 A rhetorical structure analysis of anaphora in Chinese**

According to Tao (1996), who follows the notion of discourse sequence investigated by Fox (1987), discussed in Section 4.2.1. above, a full NP is generally used as first mention at the start of a new discourse sequence, that is, a series of utterances or expressions which relate to a discourse topic - that is a referent recurring throughout the sequence. Subsequent mentions may use a pronoun or a zero pronoun (Tao 1996: 493). However, while he points out that the constraints of topic continuity are crucial in determining the use of zero anaphora Tao notes that the choice as between pronoun and zero pronoun may deviate from the basic pattern to follow the sequence of micro-events within the discourse. Where the event shifts, a pronoun or even a NP may be used rather than a zero, illustrated in Tao's example:

33.

- A. 它就跳到地上来  
 ta jiu tiao dao di-shang-lai,  
 it then jump arrive ground on come

到底 给它 抓住了  
 0 dao-di gei ta zhua-zhu-le  
 till-end by it catch-stop PFV

- B. 是吗  
 shi ma?  
 right Q

- A. 那蛾子飞来蹿去的，一下就到这里边  
 nei ezi fei-lai cuan-qu de, 0 yixia jiu dao zhe-bian  
 that moth fly-come-dash-go CSC sudden then arrive this side  
 来了。一下又把它抓住了  
 lai-le. Yixia 0 you ba ta zhua-zhu-le.  
 come-PFV. sudden again BA it catch-stay-PFV.

- A. 'It (the cat) jumped down. (The moth) finally was caught by it.

- B. 'Is that so?'

- A. 'That moth flew all around. In a flash (it) flew over here, all of a sudden, (it) caught it again.'

(Tao 1996: 491)

Tao claims that the reason for the zero pronoun in A's second utterance is the marked re-introduction of the cat into the discourse.

In what Tao terms the 'return pop' pattern (after Reichman, 1981 and Fox 1987) a zero pronoun may be used when a prior referent is reintroduced into the discourse after some absence (Tao 1996: 493). He notes that the use of the

return pop in spoken discourse indicates that the speaker wishes to return to a referent in a prior discourse sequence, rather than that of the immediately preceding discourse.

Tao notes that in Chinese, return pops may have specific discourse cues: the choice of a zero anaphor in Chinese discourse may only be made possible by the use of certain verbs associated with the referent. For example in Tao's example of a conversation about an anaesthesia gun, where the gun is frequently referred to by means of zero, the verbs in the discourse all cue that particular referent semantically eg: design, shoot, inject, filled with (Tao 1996: 497). Tao maintains that his data show that in Chinese conversational discourse the use of zero, a low information anaphoric device, to re-introduce a distant referent may seem counter-intuitive, but is in fact marked. 'When Chinese speakers process discourse information, their mental activities must be focused on the overt or covert cues that are associated with zero anaphors' (Tao 1996: 505). It may be that Chinese speakers pay significantly more attention to those cues than do speakers of languages which do not have abundant zero anaphors. A Chinese speaker may pay more attention to providing specific cues and a listener may listen for those cues. Other researchers suggest the importance of the verb in anaphoric choice and resolution. Tyler (1984) in a study of Dutch speaking children, found that if there were constraints in the discourse which might lead to ambiguity, children resolved the anaphora by using the protagonist appropriate for the verb, or by introducing an entirely new entity. Hu (1999), in a paper which refutes the notion of a distinction between finite and non-finite verbs in Chinese, notes the 'verbiness' of Chinese; in terms of language typology, as a non-tensed language, adjectives in Chinese are verby rather than nouny. As I will discuss later, with reference to my data, it may be that the semantic content of the predicate (the verb phrase) may be much more important in

anaphoric choice and resolution than that of the argument (noun phrase), especially where the noun phrase is so attenuated as to have a zero form.

#### **4.4 Anaphoric resolution**

While a speaker or writer makes anaphoric choice, that is, selects appropriate anaphors to use, a listener or reader must make anaphoric resolution, that is, decide which referents those anaphors refer to. There seem to be fewer studies of anaphoric resolution than of anaphoric choice. The study by Yang et al. above is related to anaphoric resolution, though its main aim is to find a centring model for anaphoric reference.

##### **4.4.1 A study of language specific discourse strategies in the processing of zero anaphors by Chinese readers and English readers**

In a study purporting to establish a difference in English and Chinese speakers' processing of zero pronouns, Tao (1993) suggests that while zero anaphora may be ambiguous in single sentences it is rarely so in discourse. His aim in this study was to investigate how L1 information processing strategies of speakers of different languages (in this case English and Chinese) influence the processing of L2. His method was to compare the use of zero anaphora in Chinese with the use of overt nouns and pronouns in English and thereby find an explanation of when and why speakers of Chinese, but not speakers of English, can comprehend discourse which contains frequent zero pronouns (Tao 1993: 241).

Tao notes that it may not always be easy to decide whether there is actually a zero pronoun in a text, since actual discourse may provide more than one slot in which to place a zero pronoun (Tao 1993: 40). He proposes that zero anaphora is used when an event is continuous and pronouns are used when there is discontinuity, that is, the zero pronoun serves to form continuity, while zero pronouns signal a shift of events and a full NP is used when other

devices are insufficient (Tao 1993: 115, 116), as discussed above in Section 4.3.2.2.8.

He suggests that anaphora should be analysed at three different levels: through a global view of discourse; shift of action or state by the same referent; continuation of an event (Tao 1993: 55). In this type of analysis, any protagonist that is reintroduced into a sequence after an absence will be manifest as a full NP, or Level 1 referent. It will then become a Level 2 referent, that is a pronoun, as events or episodes in the discourse change, and finally will become a Level 3 referent, that is a zero pronoun (Tao 1993:59). Tao thus confirms a pattern of attenuation found in the event continuity of the set of clauses that make up the topic chain (Tao 1993: 71), this pattern enabling the reader/hearer to make successful anaphoric resolution.

In his data Tao found almost as many full NP mentions as zero mentions - and in exactly the same proportions for written and spoken discourse. This high proportion of NPs appears contrary to previous topic continuity theories reviewed above, such as that of Givon (1983), Li and Thompson (1979) and Xu (1990), but Tao notes that this may be due to the information status of referents, or syntactic or discourse constraints (Tao 1993: 95).

Tao discusses the phenomenon of zero pronoun as first mention (see discussion of null topic in Section 4.2.2.3). Tao calls this 'delayed first mention' and postulates that the reader/hearer may recognise this as new information and process it while waiting for the identity of the zero to be revealed, since certain activities will be associated with certain referents (Tao 1993: 111). The discourse role of the delayed first mention, or zero topic, would seem to be a certain kind of cataphora, perhaps used for the purposes of suspense. Tao notes that such constructions may give the reader/hearer an added cognitive burden (Tao 1993: 112).



Tao posits three environments which allow the use of zero anaphora:

- a. a single referent environment;
- b. multi-referent environment;
- c. return-pop environment.

He notes that in a discourse where multiple referents (2 or more) are coded by zero pronouns, there is potential ambiguity, but Chinese native speakers would rarely find this a problem, as they probably pay attention to the discourse cues. Such cues, as discussed briefly above in Section 4.3.2.2.8, include information in the previous context, semantic information of verbs and language users' general knowledge of the world (Tao 1993: 177). These factors are most clearly seen at work in Tao's example of the moth and the cat (quoted above in Section 4.3.2.2.8) where the information provided by the meaning of the verb combines with the language user's general knowledge of the world to interpret the zero pronoun - a cat catches a moth, not vice versa.

Tao claims that while English speakers may focus on the syntactic structures before they can get sufficient semantic information out of a text, Chinese speakers may focus on the semantic features which are not marked syntactically, but are none the less carried by verbs (Tao 1993: 293). The study provides an excellent and detailed description of factors involved in anaphoric resolution for the Chinese native reader/listener, but exhibits certain weaknesses. As with one or two of the experimental studies above, conditions for the two sets of native speakers differed markedly. Tao's English native speakers were monolingual undergraduates and were tested using texts that 'did not look like natural English' (Tao 1993: 253). The Chinese native speaker subjects, however, were postgraduate bilinguals and were tested on similar English texts and natural Chinese texts. The differences found between the two groups of subjects were in their own ratings of their confidence in anaphoric resolution of the texts, and their

ratings of comprehensibility of the texts. If a subject rates a text as having low comprehensibility, it does not necessarily mean he does not understand it - he may simply be exercising analytical faculties. Subjects in both Tao's groups were required to time themselves and in all tests the Chinese native speaker subjects took appreciably longer to respond. There is no doubt that the study highlights the factors in anaphoric resolution in Chinese, and shows differences between English and Chinese structures which are already well known. However, I feel that it cannot be said to show any objective evidence of language specific discourse strategies owing to the deficiencies noted above. However, Tao's theoretical model provides strong support for discourse based strategies of anaphoric resolution of the zero pronoun in Chinese.

#### **4.4.2 A study of anaphoric resolution in Chinese speaking children and adults**

Wong (1990) carried out experimental investigations into the way in which Chinese (Cantonese) speaking children interpreted intra-sentential anaphora and what criteria they would use in interpreting cross-sentential anaphora. She looked at possible differences in anaphoric resolution as between reading and listening, and whether children attended to orthographic information in the pronoun.

Wong notes that, as discussed by writers such as Marlin-Wilson and Tyler (1980), children may be sensitive to different sources of information from adults in the use and resolution of anaphora, though researchers do not always agree on findings relating to the development of anaphora. Some (eg Lust et al., 1980, 1981, 1983) have found a directionality constraint: in a right branching language, such as English, anaphora will be constrained rightward and in a left branching language such as Korean, Japanese or Chinese, anaphora will be constrained leftward. Lust found that English speaking children aged four to seven often avoided backwards anaphora (cataphora),

basing their choice on a hypothesis that in English an anaphor should follow its referent. However, O'Grady et al. (1986) found that Korean and Japanese speaking children did not use any directionality constraint, a factor which led them to the conclusion that directionality preference for anaphora would be the same for all languages, and that linearity was the critical factor. Thus backward patterns of co-reference (cataphora) would be marked in all languages.

Wong based her experiments on the hypothesis that in Chinese, 'a most common syntactic position of a null pronoun is the subject position of the second co-ordinately conjoined clause' (Wong 1992). This supports Chen's claim that the most persistent argument (the topic) is likely to occupy pre-verbal position. It is also a notion that is in line with syntactic approaches to the reference chain, and at the same time reflects Li and Thompson's (1979) and Chen's (1986) conclusion that conjoinability of clauses predicts the occurrence of null subjects.

Wong studied primary school children from a working class community and adults in their mid-twenties who had completed secondary education. All her subjects were Cantonese speakers and she notes that in the Hong Kong system they would be reading and writing Modern Standard Chinese (see Chapter One), but would probably be reading aloud in class using Cantonese phonology. She notes that the syntax of the sentences used in her study 'is virtually the same in both Cantonese and Modern Standard Chinese' (Wong 1992).

Wong concluded, by means of comparison of adults and children, and children of various ages, that while some types of anaphora are acquired early, others are learned from language experience. She found that factors affecting anaphoric resolution included the type of anaphora (whether null or

pronominal), the linguistic context i.e. the degree and type of embedding, whether or not there was a conjoining conjunction, and the distance between the antecedent and the anaphor. Anaphoric resolution differed in listening and reading and was affected by the amount of 'learning' the children already possessed. Wong found that there was some evidence of developmental differences in complex sentence structures. Most significantly, she found that when faced with highly complex sentences and deep embedding, children adopted a surface structure strategy. For example, in her test item as follows:

$Np_i$  saw  $Np_j$  PRO ask the teacher whether 0 could go to the playground.

Question for C: Who asked the teacher?

Question for D: Who wanted to go to the playground?

Wong notes that in responding to the C question, children tended to make the same preferred anaphoric resolution as adults, but in responding to the D question, showed developmental differences. Grade 2 children (the youngest in the study, average 8 years) seemed to be not secure when listening to this sentence structure, but seemed to be more secure when reading similar test items. Wong suggests that in reading, they can 'look back,' an option which is not available in listening. She further notes that in the resolution of functional anaphora both children and adults were affected by whether the anaphor was zero or pronominal. In Wong's examples as follows:

E item:

$NP_i$  hit  $NP_j$  then 0 immediately went away.

Question: Who went away?

N item:

$NP_i$  hit  $NP_j$ . He then immediately went away.

Question: Who went away?

Wong found that both children and adults had a preferred choice in E but children found N ambiguous and adults showed a less preferred choice. She

notes that it is likely that temporal conjunctions such as BIAN 便 (then) used in these examples, which express high conjoinability, had not been grasped by Grade 2 children, so they were unable to apply this grammatical knowledge to anaphoric resolution. However, she claims that lexical knowledge was applied by both children and adults in anaphoric resolution, as in the following example:

NP<sub>i</sub> drew NP<sub>j</sub> (eg 'a fairy'), 0 very happy.

Question: Who was very happy?

Wong notes that there was no difference in group means as between listening and reading, and that children's preference was similar to that of adults, ie the preferred anaphoric resolution was NP<sub>i</sub>. She concludes that this could be attributed to selection restrictions of the adjectival phrase 'happy,' in other words, humans are more likely to be happy than fairies! Given that this experimental item offered no context, and children were unable to base their resolution on topicality, I think it is more likely that children identified with the human referent, that is they applied a schema from their own experience, a notion to which I will return in discussion of my data. Another possibility is that their resolution was affected by a criterion something like the animacy criterion - perhaps in a hierarchical discourse system null arguments are more likely to refer back to natural beings than to supernatural beings.

Wong notes that children in her study did not make use of orthographic information (ie about gender) in anaphoric resolution. She concludes that this is due to lack of experience in reading.

#### **4.5 Summary and conclusion**

Chinese may be considered a pro-drop language, but by contrast with European pro-drop languages it does not have a rich pattern of

morphological inflection to aid in anaphoric choice or anaphoric resolution. Chomsky's empty category theory requires recoverability in the context, which is possible with richly inflected languages, but less so with Chinese. Lu (1986), supported by Xu (1991) and Zhang (1996), however, maintains that any deletion, or ellipsis, in Chinese requires transposition and is marked. Chinese has a tendency to zero pronominal reference, but also to verblessness, and to the omission of connectives.

Anaphoric reference usually follows a pattern of continuity from NP through pronoun to zero. In discourse, this is reflected by the topic chain, in which the most continuous, accessible topic is more attenuated in form, that is a pronoun, or a zero pronoun, and the least continuous, inaccessible topic is less attenuated in form, that is a NP. This discourse based description may be linear, as in the view of Givon (1983) or hierarchical as in the view of Chafe (1980), Fox (1987), Chen (1986) and Tao (1996). Some writers regard the topic comment structure in Chinese as a syntactic structure (Shi 1991). A number of factors have been found to affect both anaphoric choice and anaphoric resolution: for example unexpectedness and contrastiveness may require a pronoun or NP, while conjoinability may require a zero pronoun (Li and Thompson 1979); clause boundaries, episode boundaries and empathy may require NP or pronoun while intra-clause or sentence anaphora may require zero pronoun (Clancy 1980). Some writers claim that proximity of antecedent is a criterion for choosing a zero pronoun (Reinhart 1976, 1983; Xu 1990), but this may simply be a default solution (Chen 1986). Chen (1986) claims that high noteworthiness and persistence are criteria in choice of zero anaphora. Huang (1994) takes a Gricean view of anaphora, suggesting that anaphoric choice is based on Levinson's principles of providing sufficient information in an appropriate manner. Tao (1996) suggests that the pattern of anaphora follows the sequence of micro-events in discourse and that zero anaphora may be used to reintroduce a previous topic, or to introduce a new topic into

the discourse. He maintains that readers/hearers of Chinese use cues such as the previous context, the meaning of verbs and general knowledge to interpret zero anaphors, and that these strategies may be language specific. Wong (1992) found that in anaphoric resolution, children showed developmental differences, acquiring some types of anaphora early, but learning other types from language experience. With certain types of complex, highly embedded sentences, children used surface structure evidence such as proximity. Wong found that both children and adults used lexical knowledge more than syntactic knowledge in anaphoric resolution, and perhaps criteria of animacy or being human (Wong 1992).

The studies reviewed above show that a discourse analysis approach may be more suitable for anaphoric choice and resolution in Chinese, and should take into account a number of criteria such as continuity, accessibility, persistence, conjoinability, the meaning of the verb, animacy and the reader/listener's own general knowledge of the world in the case of zero anaphora, and discontinuity, inaccessibility, decay, contrastiveness, unexpectedness, and new episode in the case of pronoun or NP.



## **Chapter 5: Analysis of the data: interpretations of zero anaphora in five short texts**

### **5.0 Introduction**

This chapter describes the findings of a study of parallel translations, analysed in order to discover what factors may affect bilingual readers' anaphoric resolution of the zero pronoun in Chinese. The study covers only anaphoric null arguments, so the term used henceforward is zero pronoun. The factors focused on are those suggested by the research reviewed in Chapters Two, Three and Four. These include antecedent proximity or distance, topic persistence, topic continuity, and information in the context, all of which appear to have a major positive effect on anaphoric resolution. Abstraction appears to have a major negative effect. Multiplicity of topics, the meaning of the verb, readers' schemata, return pop and noteworthiness appear to have a small but mainly negative effect on a minority of readers. Gender, number, animacy and subject/object roles do not seem to have any significant effect. The factors of antecedent distance and topic persistence are discussed in the light of short term memory, eye movements in reading, idea units and mental spaces.

### **5.1 The subjects**

Four short texts (approximately 150 characters each) were translated by a group of 117 readers, as part of an 'A' Level examination. Owing to the confidential nature of the examination, no information about the candidates is available, but all were attending secondary schools, or in a few cases FE colleges, in the British education system at the time of the examination. Questionnaire answers from a previous study of a similar examination population indicate that candidates who take this examination may have been born in the UK or may have been at school in the UK for as little as three or

four years.<sup>16</sup> In the case of this examination ('A' Level), however, questionnaires are problematic because of the confidentiality of the examination - enquiries can only be sent to head teachers, after results are published, and candidates may have left school by that time. Given the profile of examination centres (schools) it can be assumed that subjects in this study were probably between 16 and 19 years old, though a tiny minority may have been adult students at FE colleges. The spelling of names, and candidates' choice as to whether to use full or simplified characters gives a little indication of their language background: those who used full characters may have had family origins in Hong Kong or Taiwan, and those who used simplified characters may have had family origins in Singapore or the mainland PRC. Those candidates who grow up in the UK may have attended Chinese weekend schools but will have grown up in a largely English speaking environment. The Cantonese speakers sometimes use Cantonese idioms in writing in other parts of the examination. However, none of this information is sufficient in quantity or quality to be used in any statistical analysis. The examination was bilingual: as all the instructions were in English, and there were texts for translation from English to Chinese and from Chinese to English, ability in both languages was necessary to attempt the examination, though the stated aim of the examination was to test comprehension and production of Chinese.

## 5.2 The study

The data used in this study consist of parallel translations made in the course of the 'A' level Chinese examination<sup>17</sup> set by Edexcel Foundation until June

---

<sup>16</sup> With the permission of the examination board questionnaires were sent out to all candidates in one examiner's allocation in 1998, after publication of results. Out of 140 candidates only 20 replied. Some head teachers replied that pupils had left school and could not be contacted.

<sup>17</sup> The A Level examination has now ceased to exist in this form. The new examination which has replaced it is modular, consisting of AS Level and A level. The new AS and A Level are mainly target language examinations, but a small amount of non-literary translation is retained. As of June 2001, texts have titles, and glosses are used where deemed necessary. Subject matter in the translation texts

2000. The examination consisted of two papers. Paper 1 had three compulsory texts for translation from Chinese into English and a choice of three history questions, which could be answered either in Chinese or in English and which did not form part of this study. Paper 2, which did not form part of the study, consisted of two texts for translation from English into Chinese, a text of classical Chinese for paraphrase into modern Chinese and an essay.

The aim of the three translation texts set in Paper 1 was to test comprehension of Chinese, and the papers were marked to this end. The three texts set varied in style and content, as specified below. The 2000 syllabus specification for Paper 1, Questions 1, 2 and 3, (the texts used in this study) states

*'The passages will be in different styles, eg the standard prose essay style used in all parts of the Chinese speaking world, the practical descriptive style of the People's Republic of China, the literary style found in contemporary novels etc. ....The intention of this part of the paper is to assess candidates' understanding of written Chinese, and their ability to convey the meaning accurately into acceptable English' (Edexcel 2000).*

The texts were selected randomly by the setters of the papers, within the scope of the three styles. Whilst any discussion of a target audience is irrelevant, as candidates did not know who was marking their papers, beyond knowing that the examiner was probably a teacher of Chinese, they did have a specific aim, ie, to show their understanding of the texts. Elegant English was not a requirement - candidates' translations were marked on the basis of comprehension of the Chinese in comprehensible English.

For the reasons explained above, no information about the candidates is available. It is therefore not possible to do any statistical analysis which

might take into account age, sex, language background and ability, length of schooling and so on. Until now, the marking has been of translations of randomly selected texts on a relatively subjective basis, and it would not be appropriate to look for relationships between my tentative analysis of candidates' interpretations and examination scores. In future, under the new system of selection of texts, setting and marking, which conform to rigorous standards, it would be feasible to calculate correlations between normative and non-normative translations, and marks for individual questions and marks for papers overall. This might provide an opportunity to replicate, for example, Kallqvist's (1998) comparison of translations and essays and carry out comparisons of translations and comprehension texts.

From this point candidates will be referred to as subjects.

### **5.3 Procedures**

The study comprised four texts, three taken from the 1998 examination and one taken from the 2000 examination. Subjects' translations were compared for solutions to the zero pronouns in each text. It was found that there was wide variation in anaphoric resolution. Solutions fell into three broad bands - optimal solution, proximal solution and non-solution. The optimal solution tended to be uniform and normative but the non-solutions showed great variation, as shown in the charts in Appendix 1.

In addition to types of solutions, subjects' continuity of reference, that is, the link between anaphor and antecedent, was plotted in two bands - plus continuity [+continuity], which was sub-divided into either singular and plural, or topic and non-topic, and minus continuity [-continuity]. For example, if a subject interpreted a zero anaphor as 'style' and continued with a singular pronoun, that would be +continuity singular. If a subject interpreted the same zero anaphor as 'styles' and continued with a plural

pronoun, that would be +continuity plural. Switching from a singular pronoun to a plural pronoun would be -continuity.

Initially an analysis was made of each individual solution to each zero pronoun and pronoun or proform in the four texts. These more detailed analyses are shown in Appendix I. The range of solutions given shows how individual readers vary in their comprehension. However it was necessary to order the data in some way so as to achieve a workable banding which might eventually be applied to marking.

Individual solutions were therefore grouped together. In Text 1, for example, the optimal solution of ZP 1 (zero pronoun 1) was taken to be 'he' according to raters' judgements (see Section 5.4 below). Proximal solutions and non-solutions were also grouped according to raters' judgements.

My original plan was to include only Texts 1, 3 and 4 in the analysis, and these are the texts for which raters' judgements were used. My reason for excluding Text 2 was that during the marking of the examination papers it seemed that subjects' solutions were uniform - that is to say most readers got the same answer. However, on closer inspection, at a much later date, it was apparent that there was a considerable range of solutions to text 2, even though the text was well comprehended, and most of the solutions could be regarded as optimal. I therefore decided to include this text for analysis. In addition, it was apparent that the structure of the text was a major factor in ease of comprehension: the text has only two arguments, one animate, one inanimate, and only one theme, which relates to the highly tangible business of eating. Text 2 therefore contrasts strongly with Texts 1, 3 and 4, which have multiple arguments, and abstract themes.

The absence of raters' judgements for Text 2 is a problem, but I have based my judgements on those used for the other three texts, and my own intuition.

#### **5.4 Raters**

Representative types of readers' solutions were presented to a team of seven experienced adult raters (all teachers of Chinese in secondary, further or higher education) for accuracy judgements. All of them had experience in teaching Chinese and marking translation.

Ratings were on a scale of 1-5, as follows:

5 = accurate

4 = fairly close

3 = not entirely wrong, but could be closer

2 = inaccurate

1 = inaccurate and irrelevant

Raters' judgements of the acceptability of the translations varied very slightly, but in all cases the most popular solution of the subjects was consistent with the solution regarded as most accurate (5 on the scale) by at least six out of seven of the raters, and taken to be the optimal solution. Proximal solutions were those judged to be 4 or 3, non-solutions were those judged to be 2 or 1. The solutions which fell into the non-solution category would normally be regarded as errors, account for most of the variation and present interesting data at all levels of the interactive reading process. Some of these are dealt with in Section 5.7.

#### **5.5 The texts**

The texts are approximately 150 characters in length and are authentic excerpts. For the purposes of the examination and the purposes of this study, the texts were regarded as complete, and first mention in the text is regarded

as first mention overall. As presented in the examination they do not have titles, glosses are not given and proper names are not underlined. Subjects were not allowed to use dictionaries.

The initial inspiration for this study was a text set in the Edexcel A Level examination in June 1997. An analysis of this text showed a wide variety of interpretations by examination candidates. It was evident that some candidates were confused about who did what to whom, probably because they had misunderstood the reference patterns within the text (Waggot 1997). For the purpose of this thesis I undertook a more thorough study using texts from the same examination in three subsequent years. The texts were selected on the basis of contrasting reference patterns and candidates' response to them. Text 1 in this study, taken from the 1998 examination paper, has few zero pronoun references, but is complex in terms of abstraction, and candidates produced a wide variety of responses. Text 2, from the 1998 paper, has long complex reference chains, but is simple in terms of argument and theme. Subjects obviously understood it much better, but still showed some variation in response. Text 3, from the 1998 paper, differs from the others in that it has a number of inanimate arguments and a high degree of abstraction, and candidates produced a wide range of responses. Text 4, from the 2000 paper, whilst free of abstraction, and having only short reference chains, produces complexities for the reader in terms of gender and number, and candidates showed an unexpectedly wide range of responses to what should have been a relatively easy text. Data from the 1999 examination was rejected on the basis of insufficient information.

It seemed likely that these strongly contrasting texts and responses would yield a pattern of translation solutions which in future might be applied to marking.



### 5.5.0 Terms used in the analysis

The term topic refers to an argument which has a first mention (usually as NP) and a series of subsequent mentions as NP, pronoun (PRO)<sup>18</sup> or zero pronoun (ZP) and (0). Single mention arguments without further reference are not discussed except where they have affected subjects' anaphoric resolution.

With regard to the texts, persistence is measured by position in the topic chain, eg, ZP 1.1 is 3, as it is the third mention of the topic, Mr. Lu Xun, ie. the third link in the topic chain. Antecedent distance is measured as the number of zero anaphors after the last NP or pronoun mention, so for example, ZP 1.1 has an antecedent distance of 1 as it is the first zero pronoun after its co-referent NP and co-referent pronoun.

Abstract and animate refer to features of the referent and its anaphors in the text.

Subject refers to the role in the sentence of the anaphor in question, as defined by Tao (1996) - the NP which is one of the main arguments, usually the agent, of a transitive verb, or the only argument of an intransitive verb (Tao 1996).

Gender refers to the gender of the referent as presented lexically elsewhere in the text. Number also refers to singular/plural distinctions of referents made by the use of number or measure words or a singular or plural pronoun.

chengyu refers to a zero pronoun which has a *chengyu* (four character idiom or metaphor) as predicate, as described in Section 2.4.2.1.

---

<sup>18</sup> PRO as used in this study is an abbreviation for 'pronoun' and does not relate to generativist terms.

With regard to subjects' anaphoric resolution, the term optimal solution refers to the interpretation of a zero pronoun which is correct according to raters' judgements, that is, 5 on the rating scale. The term proximal solution refers to the next best solution, where the subject has brought some strategic competence to bear and which raters judged as 4 or 3 on the scale. The term non-solution refers to interpretations which would be regarded as wrong - they bear no relevance to the source text and are not part of a coherence pattern in the subject's translation, and were judged 2 or 1 on the rating scale. Omission signifies that subjects omitted either an argument and its associated VP or omitted a whole clause.

Subjects' continuity of reference (+ continuity singular or + continuity plural) refers to whether the subject has continued the reference to the first mention topic coherently, for example, as singular pronoun, or as a plural pronoun. Discontinuity of reference (- continuity) would involve a switch mid-sentence from, for example, singular to plural, not warranted by information in the text. Omissions (above) are included in discontinuity.

Syntactic relationships are represented in the diagrams as follows:

>>>>> = modifying or possessive relationship

..... = subject object relationship

| = anaphoric reference

----- = noun complement

ZP = zero pronoun (eg. ZP 1.2 is the second zero pronoun in Text 1.)

PRO = pronoun (eg. PRO 1.1 is the first pronoun in Text 1.)

NP = noun phrase

**Note on placing of zero pronouns**

Tao (1993) notes that it is not always easy to decide where a zero pronoun should be placed, depending on the nature of the predicate, particularly in the case of objects. Where arguments are not explicitly stated it is necessary to hypothesise a suitable structure, for example, a straightforward SVO structure, or a BA construction, or possibly a passive. The placing of the zero pronouns in the texts may be before or after the appropriate predicate.

### 5.5.1 Text 1

一些摹拟鲁迅先生杂文的人，往往忘却了这些杂文产生的时间和空间，目的和对象。在他的杂文中，有些地方(ZP 1.1)不得已的咬文嚼字，拐弯抹角，引经据典，声东击西，文白混用，或(ZP 1.2)故意含糊其辞，都是当时的年代，环境的产物，连先生自己也认为(ZP 1.3)诟病的。如今我们不能盲目地效仿(ZP 1.4)，(ZP 1.5)应该只是在(ZP 1.6)需要(ZP 1.7)的时候才能够用一用它。我们所主要该学的，却是那及时磨练武器，精通舞艺，以及临敌无畏的精神。

#### Suggested Translation:

A few people who emulate Mr. Lu Xun's essays always forget the time, place and target readers of those essays. In his essays, there are a few places where he (ZP 1.1) could not help 'talking pedantically,' 'beating about the bush,' 'quoting the classics,' 'placing decoys,' 'mixing ancient and modern language,' or even (ZP 1.2) deliberately disguising the meaning. This was a product of that era and environment and even Mr. Lu Xun himself thought it (ZP 1.3) was a fault. Nowadays we should not blindly copy it (ZP 1.4), but we (ZP 1.5) should just use it when we (ZP 1.6) need it (ZP 1.7). What we should learn is the spirit of tempering our weapons in good time, being practised in the art of combat and facing the enemy fearlessly.



The arguments in the text include inanimate (essays), animate human ('people,' 'Lu Xun,' 'we'), singular and plural, concrete ('essays') and abstract (the series of *chengyu*). There are four topic, or reference chains, referring respectively to 'Lu Xun, 'essays', the series of '*chengyu*' (style of writing) and 'we.'

Topic chains in this text are typical of the progression of given and new in a Chinese text, in which the object of one sentence or clause may become the subject or topic of another, or the predicate of one clause may become the argument of another (Zhao 2000), without the need for any overt marking of the change, such as a pronoun. In all cases subject zero pronouns are animate and in all cases object zero pronouns are inanimate.

### 5.5.1.2 Subjects' anaphoric resolution of zero pronouns in Text 1

|                    | ZP 1.1 | ZP 1.2 | ZP 1.3 | ZP 1.4 | ZP 1.5    | ZP 1.6 | ZP 1.7    |
|--------------------|--------|--------|--------|--------|-----------|--------|-----------|
| anteced. distance  | 1      | 2      | 1      | 2      | 1         | 3      | 2         |
| persistence        | 3      | 4      | 2      | 3      | 2         | 4      | 3         |
| abstract           | -      | -      | +      | +      | -         | +      | -         |
| animacy            | +      | +      | -      | -      | +         | -      | +         |
| gender             | M      | M      | N      | N      | not known | N      | not known |
| number             | sing.  | sing.  | sing.  | sing.  | plural    | sing   | plural    |
| subject            | +      | +      | -      | -      | +         | -      | +         |
| <i>chengyu</i>     | +      | +      | -      | -      | -         | -      | -         |
| %optimal solution  | 25     | 31     | 52     | 72     | 78        | 82     | 88        |
| %proximal solution | 26     | 12     | 31     | 19     | 16        | 10     | 7         |
| %non-solution      | 21     | 7      | 11     | 7      | 4         | 8      | 5         |
| %omission          | 28     | 49     | 6      | 2      | 2         | 0      | 0         |

The most striking thing about translations of the first two zero pronouns in Text 1 is the high number of omissions of argument + VP, ie. the whole proposition. This may be due to the effects of 'skipping' as discussed in Chapter Two - subjects may have not noticed these VPs or may have thought them unimportant. It is also possible that they simply were unable to translate them, but this would be inconsistent with their attempts at the rest of the text. Both are arguments having abstract predicates.



The next most striking pattern is the general tendency towards a greater number of optimal solutions as the text progresses, a less consistent trend towards fewer proximal solutions, and a general tendency towards fewer non-solutions and omissions as subjects progress through the text, possibly due to accumulation of contextual information. The inconsistency in the upward trend of optimal solutions for ZP 1.7 is probably due to abstraction.

Animacy, gender and number do not seem to have any effect on comprehension.

5.5.1. 3 Subjects’ continuity of reference in Text 1

|                        | ZP 1.1-ZP 1.2 | ZP 1.2 -NP1 | <i>chengyu</i> - ZP 1.3 | <i>chengyu</i> PRO 1.1 - | 'we' topic chain |
|------------------------|---------------|-------------|-------------------------|--------------------------|------------------|
| % +continuity singular | 36            | 34          | 15                      | 41                       | 0                |
| % +continuity plural   | 31            | 1           | 40                      | 29                       | 85               |
| % - continuity         | 33            | 65          | 45                      | 30                       | 15               |

The high degree of discontinuity in ZP 1.2 - NP 1 may be due to polite reference to Lu Xun as 'Mr' 先生 (*xiansheng*) without the name. This may be a cultural effect, or it may be due to the fact that *xiansheng* also means 'teacher.' The high degree of discontinuity in the *chengyu* topic chain may be due to the abstract nature of the topic (writing style expressed at first mention as a series of predicates). The 'we' topic chain shows a high degree of continuity.

### 5.5.2 Text 2

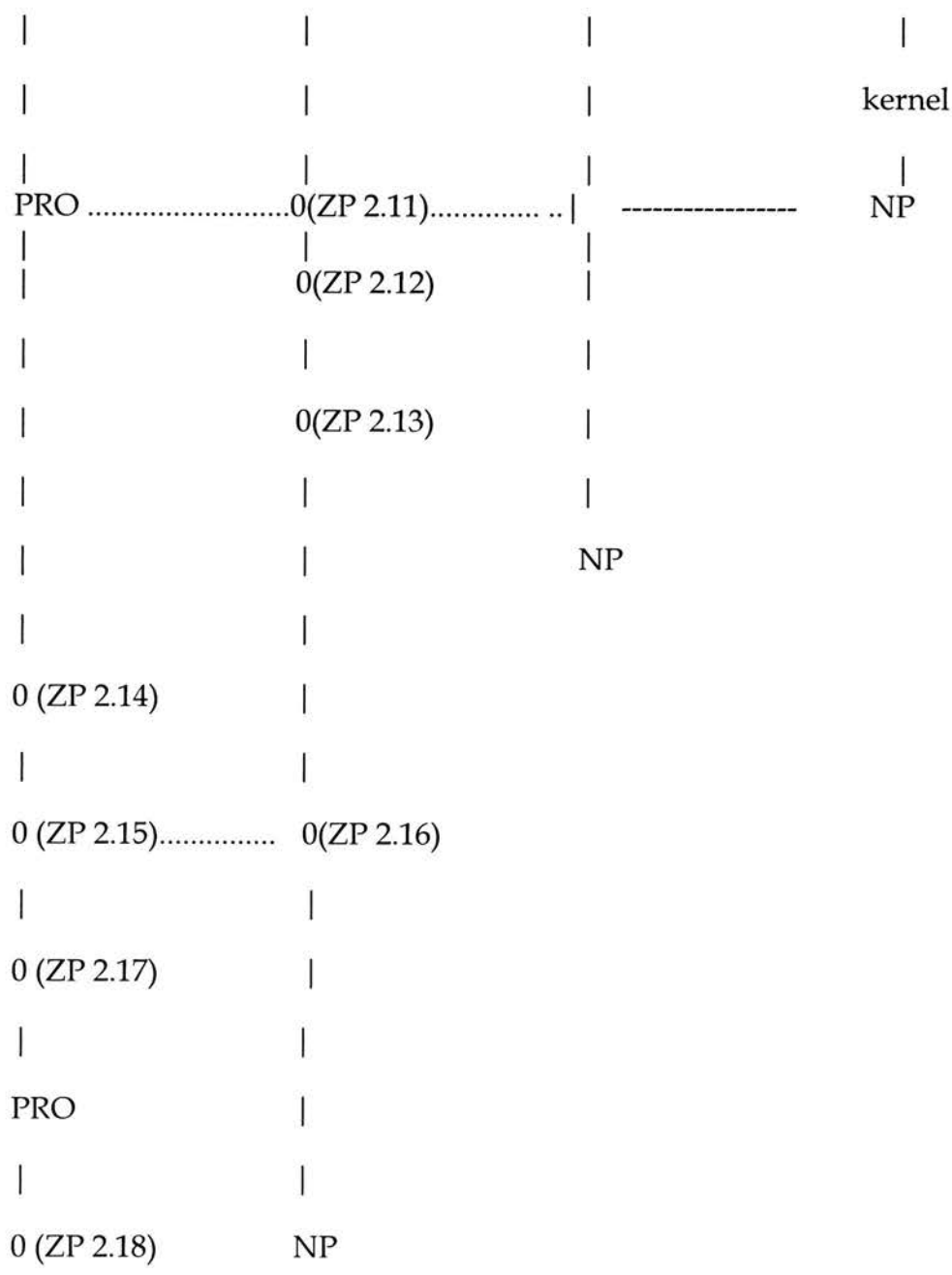
我选了一粒瓜子，(ZP 2.1)/(ZP 2.2)送交臼齿去咬(ZP 2.3)。不幸这瓜子太燥，我用力又太猛，(ZP 2.4)“格”的一响，(ZP 2.5)玉石不分，(ZP 2.6)咬成无数的碎块，我只得把黏著唾液的碎块吐在手心里，(ZP 2.7)用心挑选(ZP 2.8)，(ZP 2.9)剔出壳的碎块，然后(ZP 2.10)用舌尖舐食瓜仁的碎块。然而这挑选颇不容易，因为壳的碎块的一面也是白色的，与瓜仁无异，我误认为(ZP 2.11)全是瓜仁而舐进(ZP 2.12)口中去嚼(ZP 2.13)，其味虽非嚼蜡，却等于嚼砂。壳的碎片紧紧地嵌进牙缝里，(ZP 2.14)找不到牙签就(ZP 2.15)无法取出(ZP 2.16)。 (ZP 2.17)碰到这种钉子的时候，我就下决心，从此(ZP 2.18)戒绝瓜子。

### Suggested translation

I chose a melon seed and (ZP 2.1) put it (ZP 2.2) in my mouth to chew (ZP 2.3). Unfortunately the melon seed was too brittle, I exerted too much force and it (ZP 2.4) went 'crack'. The shell and the kernel (ZP 2.5) were indistinguishable, (ZP 2.6) bitten into countless fragments. I just had to spit out the bits, all mixed with saliva, into the palm of my hand, (ZP 2.7) carefully select them (ZP 2.8), (ZP 2.9) pick out the shell fragments, then (ZP 2.10) lick up the fragments of the kernel with my tongue. However, this selection was extremely difficult, because one side of the shell fragments was white, the same as the kernel. I mistakenly thought it (ZP 2.11) was all kernel and licked it (ZP 2.12) into my mouth to chew (ZP 2.13). It was horrible. The shell fragments got stuck in between my teeth and if I (ZP 2.14) hadn't found a toothpick I (ZP 2.15) wouldn't have been able to get them (ZP 2.16) out. (ZP 2.17) Having run into this snag, I made a vow never (ZP 2.18) to eat melon seeds again.

5.5.2.1 Reference chains and topic continuity in Text 2

|                     |                  |                           |
|---------------------|------------------|---------------------------|
| I (narrator ) ..... | melon seed       |                           |
|                     |                  |                           |
| 0 (ZP 2.1).....     | 0 (ZP 2.2)       |                           |
|                     |                  |                           |
|                     |                  |                           |
|                     | 0 (ZP 2.3)       |                           |
|                     |                  |                           |
|                     | NP               |                           |
|                     |                  |                           |
| PRO                 | 0 (ZP 2.4)       |                           |
|                     |                  |                           |
|                     | 0 (ZP 2.5)       |                           |
|                     |                  |                           |
|                     | 0 (ZP 2.6)       |                           |
|                     |                  |                           |
| PRO .....           | .....            | fragments                 |
|                     |                  |                           |
| 0 (ZP 2.7) .....    | .....            | 0 (ZP 2.8)                |
|                     |                  |                           |
| 0 (ZP 2.9).....     | .....            | NP(shell fragments)       |
|                     |                  |                           |
| 0 (ZP 2.10).....    | .....tongue..... | .....NP(kernel fragments) |
|                     |                  |                           |
|                     |                  | NP (shell fragments)      |



Note: *quan* (all/completely) is taken to be adverbial, and referring to the melon seed as a whole.

There is only one animate, human topic which is 'I' the narrator and four inanimate topics, a melon seed, fragments of the melon seed, fragments of the melon seed shell, the kernel of the melon seed. (Fragments of the kernel have only a single mention.) The fragments and the kernel are both in a hyponymic relationship to the melon seed. The animate topic chain is very long, having 13 points, five of which are pronouns, including the first mention.

### 5.5.2.2 Subjects' anaphoric resolution of zero pronouns in Text 2

|                    | ZP 2.1    | ZP 2.2 | ZP 2.3 | ZP 2.4 | ZP 2.5 | ZP 2.6 | ZP 2.7    | ZP 2.8 | ZP 2.9    |
|--------------------|-----------|--------|--------|--------|--------|--------|-----------|--------|-----------|
| anteced. distance  | 1         | 1      | 2      | 1      | 2      | 3      | 1         | 1      | 2         |
| persistence        | 2         | 2      | 3      | 5      | 6      | 7      | 5         | 2      | 6         |
| subject            | +         | -      | -      | +      | +      | +      | +         | -      | +         |
| abstract           | -         | -      | -      | +      | -      | -      | -         | -      | -         |
| gender             | not known | N      | N      | N      | N      | N      | not known | N      | not known |
| number             | sing      | sing   | sing   | sing   | sing   | sing   | sing      | plural | sing      |
| animacy            | +         | -      | -      | -      | -      | -      | +         | -      | +         |
| <i>chengyu</i>     | -         | -      | -      | -      | +      | -      | -         | -      | -         |
| %optimal solution  | 100       | 97     | 86     | 92     | 44     | 90     | 93        | 90     | 91        |
| %proximal solution | 0         | 0      | 3      | 0      | 12     | 10     | 0         | 0      | 0         |
| %non-solution      | 0         | 0      | 0      | 0      | 8      | 0      | 3.5       | 5      | 5         |
| %omission          | 0         | 3      | 11     | 8      | 36     | 0      | 3.5       | 5      | 4         |

Table for Text 2 continued:

|                       | ZP<br>2.10   | ZP<br>2.11 | ZP<br>2.12 | ZP<br>2.13 | ZP<br>2.14   | ZP<br>2.15   | ZP<br>2.16 | ZP<br>2.17   | ZP<br>2.18   |
|-----------------------|--------------|------------|------------|------------|--------------|--------------|------------|--------------|--------------|
| anteced.<br>distance  | 3            | 4          | 5          | 6          | 1            | 2            | 7          | 3            | 1            |
| persist-<br>ence      | 7            | 8          | 9          | 10         | 9            | 10           | 11         | 11           | 13           |
| subject               | +            | -          | -          | -          | +            | +            | -          | +            | +            |
| abstract              | -            | -          | -          | -          | -            | -            | -          | -            | -            |
| gender                | not<br>known | N          | N          | N          | not<br>known | not<br>known | N          | not<br>known | not<br>known |
| number                | sing         | plural     | plural     | plural     | sing         | sing         | plural     | sing         | sing         |
| animacy               | +            | -          | -          | -          | +            | +            | -          | +            | +            |
| <i>chengyu</i>        | -            | -          | -          | -          | -            | -            | -          | -            | -            |
| %optimal<br>solution  | 92           | 89         | 81         | 66         | 85           | 85           | 99         | 91           | 89           |
| %proximal<br>solution | 0            | 0          | 0          | 0          | 9            | 10           | 0          | 0            | 7            |
| %non-<br>solution     | 3            | 5          | 4          | 3          | 3            | 2            | 0          | 6            | 3            |
| %omission             | 5            | 6          | 15         | 31         | 3            | 3            | 1          | 3            | 1            |

Text 2 showed the narrowest range in subjects’ solutions, and the closest overall agreement between subjects and raters. Differences of interpretation among subjects occurred only with a new event (the splitting of the melon seed) or introduction of a new topic (the fragments). In this case the intrusion of both new event and new topic are complicated by the use of a *chengyu* metaphor (see Section 2.2.5.1). The *chengyu* occasioned a wider range of interpretations than any other item in this text. The argument of the verb YAO, 咬 (to bite), also showed a wider range of interpretation, owing to the

verb's flexible transitivity. As an argumentless verb it could be translated as 'I bit it' or 'it was bitten.'

### 5.5.2.3 Subjects' continuity of reference in Text 2

|                       | ZP 2.2 - ZP 2.3 | ZP 2.3 - ZP 2.4 | <i>guazi</i> - ZP 2.6 | ZP 2.11 - ZP2.12 - ZP 2.13 |
|-----------------------|-----------------|-----------------|-----------------------|----------------------------|
| %+continuity singular | 98              | 86              | 69                    | 22                         |
| %+continuity plural   | 0               | 0               | 16                    | 54                         |
| %- continuity         | 2               | 14              | 15                    | 24                         |

|                     | PRO - ZP 2.14 - ZP 2.15 | ZP 2.17 - PRO - ZP 2.18 |
|---------------------|-------------------------|-------------------------|
| %+ continuity 1sg   | 52                      | 68                      |
| %+ continuity other | 35                      | 23                      |
| %- continuity       | 13                      | 9                       |

There is a tendency for the degree of continuity of reference for the melon to decrease as subjects progress through the text. The most marked decrease in continuity occurs at the point of introduction of a new argument - the fragments. This is also the point at which a new event - the splitting of the melon seed - occurs. There is also an instance of return pop - the kernel 'pops' back as an interfering referent in the fragments topic chain. The decrease in continuity may be due to the hyponymic relationship of the seed and its component parts - subjects simply referred back to the melon seed rather than fussing with its constituent parts.

Continuity of reference for the narrator (1sg) increases through the text.



### 5.5.3 Text 3

当幼树从松林的废墟上破土而出，(ZP 3.1) 在天边舞动那皎洁的苗条，这不像是(ZP 3.2)又回到了破晓的黎明，(ZP 3.3) 看到了新一代的大森林露出了它灿烂夺目的曙光吗！凭著它那纤弱 苍白的一束嫩条，幼树是多么坚毅和勇敢的先驱者，(ZP 3.4) 似乎是不自量力竟要在一片荒漠贫瘠的土地上从头拓荒，(ZP 3.5) 要重建一座比已经有过的都更加伟大的森林。它又是怎样达观和睿智的开拓者，还要荫庇和哺育那按照大自然的安排将要取代它的森林。

### Suggested translation

When the young tree breaks through the soil on waste ground in the pine forest and (ZP 3.1) waves its pure shoots on the sky line, isn't this like (ZP 3.2) returning to the dawn and (ZP 3.3) seeing a new generation of great forest showing its blazing, eye-catching glory? Depending only on its weak shoots, what a determined, brave pioneer the young tree is. It (ZP 3.4) seems to want to start afresh on poor desert soil; it (ZP 3.5) wants to build a forest even greater than before. It is an optimistic and wise pioneer, and will protect and nurture the pine forest which, according to the laws of nature, will replace it.



Text 3 contains a number of inanimate arguments, but no animate arguments except the 'pioneer' which is used metaphorically. All the pronouns which occur in the text are neuter singular. All the arguments which have a sustained topic chain are related semantically - young tree, shoots, forest, pine forest. There are no overlapping subject/object roles. All the zero anaphors in the text are subjects. All the topic chains in the text conclude with a pronoun or NP.

### 5.5.3.2 Subjects' anaphoric resolution of zero pronouns in Text 3

|                     | ZP 3.1 | ZP 3.2    | ZP 3.3    | ZP 3.4 | ZP 3.5 |
|---------------------|--------|-----------|-----------|--------|--------|
| antecedent distance | 1      | 1         | 2         | 1      | 2      |
| persistence         | 2      | 2         | 3         | 5      | 6      |
| abstract            | -      | +         | +         | -      | -      |
| animacy             | -      | -         | -         | -      | -      |
| gender              | N      | not known | not known | N      | N      |
| number              | sing   | not known | not known | sing   | sing   |
| subject             | +      | +         | +         | +      | +      |
| %optimal solution   | 45     | 44        | 27        | 48     | 85     |
| %proximal solution  | 27     | 27        | 34        | 17     | 5      |
| %non-solution       | 19     | 16        | 17        | 12     | 5      |
| %omission           | 9      | 13        | 22        | 23     | 5      |

In Text 3, according to the table there is no apparent negative effect of antecedent distance on anaphoric resolution. However, the very long 'young

tree’ topic chain (8 points) has a pronoun at point 3 and a NP at point 4, in effect reminding the reader of the topic.

There is a general trend towards a greater number of optimal solutions as the text continues and also a general trend towards fewer proximal solutions and non-solutions, suggesting that accumulated information from the text aids the reader in comprehension. However, the number of omissions is very irregular. ZP 3.3 bucks the trend. This may be to do with the nature of the verb - *kan* to see, which would normally require an animate subject, but in the text has an abstract null argument (‘this’). In this text, the null argument refers to a situation and is probably best interpreted in English as a null argument with a gerund. It could be the high level of abstraction which disrupts comprehension in this case, or the problem of dealing with a new concept - having started with trees growing on waste land, the text suddenly switches to ‘returning to the dawn.’

**5.5.3.3 Subjects’ continuity of reference in Text 3**

|                         | ZP 3.1- ZP 3.2 | ZP 3.2-ZP 3.3 | NP 3.1 -ZP 3.4-ZP 3.5 |
|-------------------------|----------------|---------------|-----------------------|
| %+continuity singular   |                |               | 68                    |
| %+ continuity plural    |                |               | 16                    |
| %+ continuity topic     | 54             | 29            |                       |
| %+ continuity non-topic | 36             | 31            |                       |
| %- continuity           | 10             | 40            | 16                    |

The continuity table shows a drop in continuity with ZP 3.2 and ZP 3.3 (the ‘this’ topic chain, suggesting that abstraction and new concepts may disrupt the reader. More subjects maintained continuity throughout the ‘young tree’

topic chain, even when they opted for proximal rather than optimal solutions. The length of the chain and the repetition of the topic as a NP midway and at the end of the chain could increase the 'aboutness,' or accessibility of a topic, thus assisting the reader in comprehension.

#### 5.5.4 Text 4

。。。。小蒋在大陆好歹是个文化圈里熏陶出来的，当天晚上酒楼里灯火辉煌，因为过年的缘故，前来欢宴的顾客如云。小蒋器宇轩昂地率领一溜儿高矮不齐的队伍踏上楼梯，(ZP 4.1)挥手叫来服务员，(ZP 4.2)选中一张圆桌 (ZP 4.3)围席入座。这是一次奇特的晚宴，坐在主人位置上的是珠环翠绕的小蒋，身旁偎著她的那个穿著珠绣绸缎中式丝棉袄的小少爷，而被宴请的主宾则是一位面色凄苦的中年妇女，(ZP 4.4)旁边依次排列著四个焉头搭脑的男孩和女孩。小蒋央求她胡搅蛮缠的宝贝赐给片刻的宁静，(ZP 4.5)忙著替另外的四个孩子夹菜盛汤，而她自己光顾滔滔不绝地说话，(ZP 4.6)顾不上吃饭。

#### Suggested translation

Anyway, Xiao Jiang was a person who was steeped in the ways of the cultural circles of mainland China. That evening the lamps were bright in the restaurant because it was new year and crowds of customers had come to celebrate. Xiao Jiang proudly led her miscellaneous troop up the stairs, (ZP 4.1) waved her hand to a waiter, (ZP 4.2) chose a table and they (ZP 4.3) all sat down around it. This was an unusual party. Sitting in the host's seat was Xiao Jiang, decked out in jewels, and snuggling up to her was her small son who wore a Chinese style silk padded jacket embroidered with pearls. The guest of honour was a pale middle aged woman. At her (ZP 4.4) side was a row of four gloomy looking boys and girls. Xiao Jiang begged her darling baby, who was pestering her, to give her a moment's peace, and (ZP 4.5) busied herself helping the other four children to soup and vegetables, while she herself talked non-stop and (ZP 4.6) was too busy to eat.

#### 5.5.4.1 Reference chains and topic continuity in Text 4

Xiao Jiang

|                  |            |                 |            |
|------------------|------------|-----------------|------------|
|                  |            |                 |            |
| NP.....troop     |            |                 |            |
|                  |            |                 |            |
| 0(ZP 4.1)        |            |                 |            |
|                  |            |                 |            |
| 0 (ZP 4.2)       |            |                 |            |
|                  |            |                 |            |
|                  | 0 (ZP 4.3) |                 |            |
| NP               |            |                 |            |
|                  |            |                 |            |
| PRO.....son      |            |                 |            |
|                  |            |                 |            |
|                  |            | guest-----woman |            |
|                  |            |                 |            |
|                  |            | 0(ZP 4.4)       |            |
|                  |            |                 | boys&girls |
| NP.....          | NP         |                 |            |
|                  |            |                 |            |
| 0(ZP 4.5).....NP |            |                 |            |
|                  |            |                 |            |
| PRO              |            |                 |            |
|                  |            |                 |            |
| 0(ZP 4.6)        |            |                 |            |



There are five topic chains in Text 4: Xiao Jiang, her ‘troop’ of guests, her son, a middle aged woman and a group of four boys and girls. Xiao Jiang is in subject role except at first mention of her son, where ‘she’ becomes object for one mention only in the topic chain. The middle aged woman remains subject over two mentions, the four boys and girls are subject at first mention and indirect object at second mention.

5.5.4.2 Subjects’ anaphoric resolution of zero anaphors in Text 4

|                    | ZP 4.1 | ZP 4.2 | ZP 4.3 | ZP 4.4 | ZP 4.5 | ZP 4.6 |
|--------------------|--------|--------|--------|--------|--------|--------|
| anteced. distance  | 1      | 2      | 1      | 1      | 1      | 1      |
| persistence        | 3      | 4      | 2      | 2      | 8      | 10     |
| subject            | +      | +      | +      | +      | +      | +      |
| abstract           | -      | -      | -      | -      | -      | -      |
| gender             | F      | F      | mixed  | F      | F      | F      |
| number             | sing   | sing   | plural | sing   | sing   | sing   |
| animacy            | +      | +      | +      | +      | +      | +      |
| %optimal solution  | 78     | 68     | 32     | 79     | 84     | 89     |
| %proximal solution | 0      | 22     | 55     | 16     | 8      | 5      |
| %non-solution      | 20     | 6      | 0      | 1      | 7      | 2      |
| %omission          | 2      | 4      | 13     | 4      | 1      | 4      |

There is a trend towards a greater number of optimal solutions as the text progresses. Those subjects who opted for a non-solution were those who assumed that Xiao Jiang was male, and did not make a suitable adjustment on encountering a feminine pronoun (occurring twice in the text). Evidence from subjects’ corrections suggest that about a third of subjects at first took Xiao

Jiang to be a man, but changed their minds on encountering the feminine pronoun. Proximal solutions tend to decrease through the text, and non-solutions are irregular. Optimal solutions fell and proximal solutions rose at ZP 4.2. Proximal solutions ('they' ie the troop of guests and 'the waiter') were feasible in terms of common sense, and may have been chosen on the basis of proximity. Optimal solutions plummet and proximal solutions soar at ZP 4.3. The verb here is crucial - 'take seats around the table' implies a group of people, rather than a single person. Proximal solutions ('she' ie Xiao Jiang) were probably based on topic continuity, rather than the semantic features of the VP.

**5.5.4.3 Subjects' continuity of reference in Text 4**

|                         | ZP 4.1 - ZP 4.2 | woman - ZP 4.4 | ZP 4.5 - ZP 4.6 |
|-------------------------|-----------------|----------------|-----------------|
| %+ continuity singular  | 60              |                |                 |
| %+ continuity plural    | 26              |                |                 |
| %+ continuity topic     |                 | 95             | 76              |
| %+ continuity non-topic |                 | 0              | 0               |
| %- continuity           | 14              | 5              | 24              |

Subjects' continuity in this text is highest for the middle aged woman topic chain, suggesting that proximity and shortness of topic chain contribute to comprehension. Continuity is high at the beginning of the Xiao Jiang topic chain, suggesting that persistence and 'aboutness' contribute to comprehension, but lower at the end of the topic chain, after the interruption of a second female person (the middle aged woman) into the narrative, that is, return pop.

## 5.6 Summary of results

Before analysing subjects' parallel translations in any detail, it is helpful to look at whether the four Chinese source texts under consideration match the patterns of anaphora outlined in the reviews in Chapter Four.

### 5.6.1 The interdependence of syntactic and discoursal constraints on anaphoric choice in the texts used in the study

Li and Thompson (1979) concluded that 'zero anaphora in Chinese is widespread and unconstrained' and that 'there are no structural properties predicting the interpretation of the referent for zero pronouns' (Li and Thompson 1979). However, it is apparent from other studies, such as those of Chen (1986), and Huang (1994), that there are both syntactic and discoursal constraints on pro-drop in Chinese.

Xu (1990), for example, notes that the speaker or writer must ensure that a zero pronoun is located close to the noun or pronoun so that sufficient information is provided for the listener or reader. Xu shows, moreover, that where there is a pause between clauses (ie in writing, a punctuation mark ) the anaphor in the second clause is likely to be a pronoun, whereas where there is no formal break between clauses, a zero anaphor is likely to be used in the second clause (Xu 1990). Tao (1996) claims that pronouns or NPs may be used when the event shifts, which is likely to be after a clause or sentence boundary.

Chen's study (1986) shows a similar tendency, in that 70% of zero anaphors occurred in the same sentence as their antecedents, while about 50% of pronoun anaphors were located across sentence boundaries. Chen claims that his study supports Clancy's view that cognitive constraints on the listener or reader limit the range of reference he or she can cope with. (Clancy found that Japanese speakers used a zero across three or fewer clauses - any greater range would require a NP (Clancy 1980).) Fox (1987), though investigating

English anaphora, also claims that in written discourse there are limitations on pronominalisation imposed by distance of the antecedent (Fox 1987). In all four Chinese source texts in this study pronouns are most often located after punctuation breaks. In these texts zeros do occur after punctuation breaks. In fact, there are a few examples of zero pronouns in situations of non-conjoinability and it may be the case that it is not the choice of pronoun which signals non-conjoinability but the use of a punctuation break. This does not support Li and Thompson's own view that zero pronouns are used in clauses that have 'high conjoinability' (Li and Thompson 1979), but these texts are a very tiny sample and may not be representative of Chinese texts as a whole. It may, however, have to do with the relationship between short term or working memory and idea units, discussed below.

Givon's (1983) linear description of the topic chain fits in with the long sustained topic chains found in these texts, for example, the first person singular topic chain in Text 2. Givon's model does not, however, allow for a hierarchy of foregrounded and backgrounded topic chains such as are found in the four texts under discussion. Chen's (1986) hierarchical model allows for more attenuated forms (zero pronouns) to occur in syntactic slots higher up the hierarchy - they are topical and central and likely to be in preverbal (ie probably subject) position. Less topical arguments are more likely to be NP or pronoun and be in post-verbal (probably object) position (Chen 1986). Chen (1987), Xu (1990), Miao (1990), and Wong (1990) all claim zero pronouns to be more common than pronouns in subject position. This is borne out by the four Chinese source texts in this study. In each, there is a long sustained central topic with many zeros mostly in subject position - 'Mr. Lu Xun' in Text 1, 'I' in Text 2, 'young tree' in Text 3 and 'Xiao Jiang' in Text 4. The longer topic chains in the texts have early and late mention and may be said to be most focused and most accessible.

According to Siu (1988) initial mention has to do with thematic importance, while concluding mention has to do with equal or secondary importance. In all four texts there are slightly shorter topic chains, containing relatively fewer zero pronouns and relatively more pronouns and NPs, which represent the theme, that is, 'style of writing' in Text 1, 'the melon seed' in Text 2 and 'forests' in Text 3. In Text 4 topic chains other than 'Xiao Jiang' all consist of two mentions only, and are all animate - thus backgrounding all protagonists in the text other than 'Xiao Jiang.' Very short topic chains in the texts consist largely of NPs and are backgrounded.

### **5.6.2 Subjects' anaphoric resolution shown in explicitation of zero pronouns**

It is apparent that among the bilingual subjects in the study, translating these texts, there is a wide variation in interpretation of zero pronouns, even where ambiguity would not normally be expected. The pattern of optimal, proximal and non-solutions shows under what circumstances subjects are more or less likely to make accurate anaphoric resolution.

Translating from Chinese, which favours zero pronouns, into English, which favours pronouns, may result in both obligatory and optional explicitation (see Section 3.5.2.1): some predicates which do not require explicit arguments in Chinese do require explicit arguments in English. On the other hand explicit arguments are not obligatory in English in some cases, for example when an infinitive is used.

Explicitation in the target language to some extent validates the notion of null argument or zero pronoun. Li and Thompson (1979) point out that in the case of a zero pronoun there is nothing actually missing in a Chinese text, but the fact that reader/ translators insert something - a pronoun or a NP - reveals that they must have a mental representation of an argument. Zero pronouns

are recoverable from the text (Lu 1986, Xu 1991) and have a cognitive if not a textual representation. Frequently, a subject uses a NP in English where a pronoun is sufficient, making doubly clear what he or she thinks the antecedent is. The patterns of explicitation in this study support Baker's (2001) idea that explicitation may reveal cognitive processes of translation.

In general the majority of reader/translators, and specifically the subjects in this study, translate the Chinese zero pronouns they encounter into English normatively. They also achieve a continuity of reference within their own target texts; having determined upon a certain interpretation of a zero pronoun they tend to stick with it. For example, in Text 1, most of those subjects who opted for 'style' continued with a singular pronoun, and most of those who opted for 'styles' continued with a plural pronoun.

A few subjects, however, did not achieve a normative solution and where proximal solutions, non-solutions or omissions rise dramatically, it may be because certain features of the text disrupt comprehension. Section 5.6.1 above sets out certain syntactic and discourse patterns of the four texts which follow patterns found by other researchers in other texts, which may be common to Chinese written texts in general. It is likely that common, known patterns in text will assist the reader. Patterns which are unusual or infrequent, however, may disrupt reading.

### **5.6.3 Factors affecting reading comprehension**

It should be noted that this summary is based on the four texts and the 117 subjects mentioned above, and generalisations cannot be made on this basis. However, some results do corroborate research reviewed in previous chapters.

### **5.6.3.1 Antecedent distance and proximity**

Much has been made in generativist descriptions of anaphora of the importance of the proximity of the antecedent in anaphoric resolution (eg. Reinhart 1983, Huang 1989). However, recent discourse based studies of anaphoric choice and anaphoric resolution show that proximity is less important than other factors. Miao (1996, 1999) found that proximity affected processing in sentences of three clauses, but not in sentences of two clauses (Miao 1996a, 1999). Wong (1990) found that children resorted to surface structure interpretations (ie having criteria of linearity and proximity) only when encountering highly complex, deeply embedded structures, in other words as a default strategy. Chen (1986) also regards the use of proximity as a criterion in anaphoric resolution as a default solution. Huang (1994) demonstrates with common examples (nos. 31a and 31b above) that to opt for the nearest antecedent may give a wrong solution, and this is borne out by my study.

It is apparent from the tables that antecedent distance (proximity) on its own has no discernible effect on subjects' comprehension in this study. There is no clear increase or decrease in optimal solutions based on proximity or distance of antecedents.

### **5.6.3.2 The positive effect of topic persistence**

There is a marked tendency in all the texts for optimal solutions to increase as the topic persists.

#### **5.6.3.2.1 Combined effects of topic persistence and antecedent distance**

It is worth noting that while topic chains in the texts in this study are attenuated, in Givon's (1983) term, by the use of zero pronouns, they are also reinforced by alternate use of NP or pronouns. While a topic may persist for a long stretch, zero anaphors do not persist for long stretches. In texts 1, 3, and



4, texts which have a high number of arguments of similar gender, number and animacy, the most common antecedent distance for a zero pronoun is 1, that is it comes immediately after a NP or pronoun antecedent. In texts 1, 3, and 4, antecedent distance is never more than 3. Thus the reader is constantly reminded of the antecedent. Even in Text 2, the text which has easily distinguished topics, (that is, only one animate and only one inanimate) where topics persist for as many as 13 links in a topic chain, the most common antecedent distance is 1, the majority are 1, 2, or 3, and only four out of 18 zero pronouns have an antecedent distance of more than 4. The longest antecedent distance in Text 2 is 7. It would seem, then, that persistence of the topic is aided by regular reminders or reinforcement of the antecedent, by means of a NP or pronoun, thus reducing the distance between antecedent and anaphor. From a Gricean perspective, this also ensures that the speaker or writer is giving the listener sufficient information, but not too much (Huang 1994).

This finding is not consistent with Givon's notion of a cline of attenuation (Givon 1983). However, it is consistent with Zhang and Simon's (1985) results of research into short term or working memory. As previously noted, any notion of reference tracking in texts must take account of the reader's working memory, and in these texts, at least, the writer 'reminds' the reader of the antecedent. Clancy notes that a shift of reference type at a clause boundary may have to do with memory constraints (Clancy 1980). Both Rayner and Pollatsek (1989) and Chafe (1980) have related working memory to eye movements, and Chafe has related eye movements to idea units (Chafe 1980). Whilst there is no clear picture in my data of whether it is this reinforcement of the antecedent which helps readers, there is no negative effect. It seems that while attenuation signals foregrounding and givenness, proximity to the antecedent helps to keep the topic in memory.

The nature of the texts under discussion would seem to suggest that the reinforcement of the antecedent in the topic chain is not to do with disambiguation ( for example, in Text 4, the insertion of the feminine singular pronoun increases ambiguity for many readers).

#### **5.6.3.2.2 The positive effect of persistence and continuity of topic**

Persistent topic chains, no doubt aided by regular reinforcement, as well as attenuation, provide continuity for the reader. The texts under discussion are short, and do not provide scope for investigating new events or episodes and new concepts, but the very persistent topic chains show a greater number of optimal solutions than the shorter topic chains. The continuity of reference tables show that subjects were better able to track reference continuity when faced with a more persistent topic.

#### **5.6.3.2.3 The positive effect of context and or topic persistence**

It is apparent in the tables that for each of the four texts there is a tendency (not always regular) for optimal solutions to increase as the subject progresses through the text. This may indicate that the more context the reader has, the more efficiently he or she is able to build up a schema. A concrete example of this is in Text 4, where corrections in the target texts show that a number of subjects started with the idea that Xiao Jiang was male, but adjusted their reading half way through. On the other hand, it is possible that this accumulated improvement may be due to persistence of topics.

#### **5.6.4 Theme and centrality**

Persistence also contributes to centrality, and what is central may be said to be the theme of the text. The topics which have early mention and late or final mention in a text may be said to be central and thematic ( see review of Siu (1988) in Section 2.6.3.3.2.). It is apparent that while a persistent topic may be

both attenuated and reinforced, backgrounded topics (such as 'essays' in Text 1 and the 'middle aged woman' in Text 4) mostly have NP reference, which corroborates Givon's notion of least continuous, most inaccessible. However, the notions of given and new do not seem to be reflected in the topic chains in the texts in this study. Where NPs are used throughout a topic chain (for example 'essays' in Text 1) second and third mention cannot possibly be regarded as 'new' information. However, use of the NP does have the effect of distancing, in much the same way as a prolix term of address is used to express courtesy to a stranger. Or, like pronouns used to reinforce the attenuated main topic chain, a string of NPs may be a means of reminding the reader of background themes.

#### **5.6.5 No apparent effects of animacy, gender, number, subject/object**

The notion of cue validity was discussed in Section 2.6.3.2. Cue validity is the product of how often a cue is available and how often it helps the reader to the right answer (Li Bates, Liu and McWhinney 1992). Miao (1981, 1996, 1999) studied the relative importance in comprehension of cues such as word order and animacy. Miao (1981) and Li et al. (1992) found that while animacy appeared to be a more important cue in determining the subject or agent of a sentence, animacy and word order interacted in some cases.

In this study the cue of word order was not investigated and the tables show no apparent effects of animacy. Some subjects show confusion as to gender, but this seems to be largely due to application of inappropriate schemata. Number also seems to have no significant effect on comprehension.

#### **5.6.6. Negative effect of abstraction and *chengyu***

In Texts 1, 2 and 3, there is a marked decrease in optimal solutions when a zero pronoun may be said to be abstract (for example in Text 3 'Isn't *this* like *returning.....*'). According to Chen (1986) ABSTRACT is at the bottom end of

the topicality hierarchy, so it should be the last option in anaphoric resolution and is likely to cause problems. Tyler found that when children did not know the antecedent they drew on something extraneous to the text (Tyler 1984). It may be that in this case, where the antecedent is abstract and therefore difficult to pinpoint, inexperienced readers looked for something outside the text, and something that went well with the verb 'look,' ending up with personal pronouns such as 'I,' 'you,' and 'we.' Similarly, subjects achieved fewer optimal solutions when the predicate of a zero pronoun was a metaphorical or idiomatic *chengyu*. This may have to do with abstraction, or cultural background, or simply lack of knowledge - as pointed out in Section 5.1, many of the subjects may have been born and brought up in the UK and may be less conversant with *chengyu*.

## **5.7 Details of contributory effects not included in the tables**

There is some evidence of effects which are numerically small but clear, and not necessarily included in the tables above, but may be seen in the details of solutions, shown in the charts in Appendix 2. Within the category of non-solution are a variety of miscues which show particular effects, for example, the effect of verb meaning and the effect of schema.

### **5.7.1 The negative effects of multiple arguments and topics not distinguished by gender, animacy or number**

It is apparent from the tables that Text 2, which contained few, easily distinguishable topics, in terms of animacy, and number, resulted in a very high number of optimal solutions throughout, in contrast to the other three texts which contained multiple topics and arguments, not easily distinguishable. Text 4, for example, contained no abstraction, and no *chengyu* and would probably normally be expected to be easier than, for example, Text 1. However, the mean optimal solutions were the same for

Text 4 as for Text 1, ( approximately 70%), perhaps due to multiplicity of human arguments.

### **5.7.2 Negative and positive effects of verb meaning**

It is evident that taking notice of the verbs contained in the predicates of zero arguments does have an effect on reading - positive if the verb has been correctly understood, negative if the verb has been incorrectly understood or ignored. Chen (1986) notes that selectional restrictions of the predicate may give essential clues to the referent. However Chen (1999) claims that the semantic processing of a verb depends on the positioning of participants, in other words, meaning of the verb and word order interact. Li, Liu, Bates and Mc Whinney (1992) found that omission reduced the reliability of word order cues, implying that when there is a zero pronoun to deal with, the meaning of the verb may be that much more important in processing. Generally, semantic information has been found to be more important for readers than syntactic information when reading Chinese (Miao 1981, 1996, Li et al. 1992, Chen 1992, Tao 1993). Tyler (1984) found that in the case of ambiguous anaphors, Dutch speaking children used a protagonist appropriate for the verb.

This strategy may help as long as the verb is accurately understood, but may lead to misunderstanding otherwise. In this study, it is evident that correct reading of the verb helps, as in the example in Text 4: subjects reading carefully realised that Xiao Jiang could not sit 'around' a table by herself, but the group of people could do so. However, incorrect reading of the verb can result in incorrect anaphoric resolution. For example, in Text 1, a number of the non-solutions resulted from subjects' misinterpretation of the predicate as 'he was ill' rather than ' he regarded it as a fault.' Similarly, in Text 3, the transitive verb 'to wave (something)' is not unlike the intransitive verb 'to dance' and a number of readers misconstrued the zero pronoun probably because they misconstrued the verb.

In some Chinese verbs transitivity is flexible, for example 'yao' in Text 2 - 'to bite' or 'be bitten' and subjects in the study varied in their interpretation, but most opted for topic continuity, choosing 'it was bitten' rather than 'I bit it.' This type of example may allow for greater flexibility in anaphoric resolution, but this is not always the case. The example given by Huang (1994, See Examples 31a and 31b, Section 4.3.2) is of this type - 'operate' or 'be operated upon' - but only common sense can provide the correct anaphoric resolution.

### **5.7.3. The effect of schema**

As discussed above in Chapters 1 and 2 background knowledge (content schema) and the structure of a text (formal schema) contribute to reading comprehension.

#### **5.7.3.1 The importance of formal schema**

It is noticeable from the diagrams and discussion of the topic and reference chains in Section 5.5 above, that topic hierarchy and anaphoric reference show distinct patterns and play a considerable role in text structure. As pointed out in Section 5.6 optimal solutions were higher in number for zero pronouns which were part of very persistent topic chains. This may indicate that readers who are aware of the patterns of anaphoric reference within a text are more likely to comprehend it successfully.

#### **5.7.3.2 The importance of content schema**

While the vast majority of subjects in the study fixed on the correct content schema, some clung to their own schema. The details of non-solutions show that content schema may play a very important part in construing a text. For example, 'essays' in Text 1, were interpreted as 'magazines' by one subject, who then adjusted the whole text accordingly. Another subject, on the basis of the metaphors in the text, interpreted the 'essays' as being about martial



arts. In Text 4, one subject wrote throughout as if Xiao Jiang were the waitress in the restaurant. Also in Text 4, on the basis of the word 'troop' referring to Xiao Jiang's guests, and reinforced by the visual and phonological similarity of 'Jiang' to the word for general, one subject interpreted 'Xiao Jiang' as 'the general.'

This type of miscue may affect anaphoric resolution indirectly. Some subjects who thought Xiao Jiang was a man, in spite of two feminine singular pronouns in the topic chain, and adjusted the text accordingly, also turned 'her son' into 'his girlfriend.' A further third of subjects started in this way, but corrected the gender halfway through the text. Corrections showed that some subjects realised and changed all pronouns to suit, while some of them kept her as male.

Changing the sex of Xiao Jiang's son, and representing him as 'sweetheart,' 'lover' or 'mistress' may also have been due to subjects' not knowing a rather old fashioned word for 'son.' This particular example highlights the problem of different cultural viewpoints or lack of cultural knowledge. (Snell-Hornby 1988; Lefevere 1990; Venuti 1995). Teenagers reading in the twenty first century, whose language of habitual use is English, are not necessarily aware of vocabulary used by their Chinese speaking parents and grandparents. Elaborative inferences may play a part here - where cultural knowledge is insufficient to interpret the text, the reader 'fills in the blanks.' Similarly, in Text 1, a few subjects interpreted *xiansheng* (Mr. ) as 'teachers' which made a difference to anaphoric resolution further down the line. This also may have been due to a lack of cultural knowledge.

#### **5.7.4 The effect of elaborative inferences**

Text 4 does not specify the relationship between the four children and either 'Xiao Jiang' or 'the middle aged woman.' A number of subjects made



elaborative inferences (Section 1.4.8) in the form of explicit pronouns, which allocated motherhood to one or the other.

### **5.7.5 The effect of return pop**

There are two examples of return pop, a situation in which a previous referent 'pops' back as an ambiguous pronoun or zero pronoun (see Section 4.3.2.2.8). In Text 2, the whole 'melon seed' pops back in after its constituent parts, and at this point subjects' continuity of reference decreases. In Text 4, 'the middle aged woman' pops back in pronoun form, making it difficult for readers to assign roles for the two women in the text, and though this does not involve zero pronouns it has a similar effect. Tao (1996) has said that return pop 'derails the continuous discourse topic' (Tao 1996: 487) but I suspect that it also derails the reader. It not only contradicts both linear and hierarchical descriptions of topic continuity such as those of Givon (1983), Li and Thompson (1979) Chen 1986) but also flouts the Gricean recipients' corollary (Huang 1994). Return pop certainly happens, in any language, and it may be up to the writer or speaker to avoid it.

### **5.7.6 Noteworthiness and punctuation**

The phenomenon of the use of zero pronouns immediately following punctuation breaks was mentioned in Section 5.6.1. Whereas Li and Thompson (1979) claim that pronouns rather than zero pronouns are used after punctuation breaks, to indicate non-conjoinability, it is the case in the texts in this study that zero pronouns are used in this position. This is not only the case with commas and full stops, but also with inverted commas and exclamation marks. Where these highlighting indicators occur in the texts, in all cases optimal solutions nose-dive. Where a reader might expect an explicit anaphor to indicate unexpectedness or highlighting, (Li and Thompson 1979) there is none. Clancy (1980) similarly notes a shift of reference type at an episode boundary which could be to do with memory constraints (Clancy

1980). Chen suggests that pronominal anaphors occurring across sentence boundaries in most cases signal a change of schema, as compared with the use of a zero pronoun where the same schema is continued and attributes this to noteworthiness, a description which is not unlike Li and Thompson's notions of unexpected and highlighted information (Chen 1986). But this is not what happens in the texts in the study, and reading appears to be disrupted by the zero anaphors in these positions.

Could it be that it is not the anaphor that guides the reader to noteworthiness, but the punctuation? However, in the texts under discussion, these features coincide with abstraction so there is insufficient evidence to make a decisive statement. Taking punctuation conventions as a whole, it may be possible to draw some relationship between punctuation, spacing, short term memory, saccades and demarcation between idea units (see below).

### **5.8 The problem of omissions**

Some parts of the texts in the study occasioned a large number of omissions. Without subjects' introspection or some experimental backup it is not easy to explain the phenomenon of omission. It may simply be that the subjects did not know certain vocabulary items, though this would be inconsistent with their performance throughout. It may be explained by two things: in the case of Text 1, the omissions coincide with 'abstract' and 'chengyu;' or, it may be that, as suggested by Pollard (1995) readers skip low density, high frequency function words or 'empty' characters.

### **5.9 Conclusion**

Parallel translations from Chinese into English by bilingual reader/translators show that explicitation is used to interpret Chinese zero pronouns. This may show evidence of readers' mental representations of the propositions in the text. The range of subjects' anaphoric resolution shows that the majority opt

for an interpretation which is in line with raters' judgements and probably also in line with the author's intention. Higher numbers of optimal solutions coincide with textual features such as topic persistence and continuity and progress through the text. This may indicate that focus, centrality and the accumulation of contextual information assist the reader. On the other hand, higher numbers of proximal solutions, non-solutions and omissions coincide with textual features such as abstraction, return pop and multiplicity of topics or arguments, and noteworthy or highlighted events or concepts, where a zero pronoun is used, rather than a pronoun, which would normally be expected by the reader. Inappropriate personal schemata and inadequate background knowledge on the part of the reader may also affect anaphoric resolution indirectly.

### **5.9.1 The combined effects of antecedent distance and topic persistence on short term memory and cognition**

Short term or working memory is an essential part of the reading process and is of particular importance in referent tracking, or anaphoric resolution. The reader, on the one hand, must remember antecedents and relate them appropriately to anaphors, especially if the anaphors take the form of zero as they so often do in Chinese. The writer, on the other hand, must ensure that his or her anaphoric choice provides enough information for the reader, without disturbing the flow, and this means placing antecedents not too far away from anaphors.

At the same time the writer must ensure that important topics are given early and late mention and are kept in focus - attenuated, in Givon's term - while scene setting elements and contributory participants in the text are backgrounded.

The reader must also be kept informed of where ideas, or events, continue or discontinue, or are noteworthy (Chen 1986). Thus, anaphors must be placed

appropriately within the current idea unit or after a punctuation break outside the current idea unit.

It is evident in the texts in this study that important focused topics (the main protagonists) occupy long, persistent topic chains, which have a great deal of attenuation (zero anaphors) compared with much shorter backgrounded topic chains which consist of NPs or pronouns. Thus the reader is aware of what or who is the main focus of the text. At the same time, the attenuation of the main topic chains is not a strict cline from full NP to zero. Rather, the string of zeros is reinforced at regular intervals by a co-referent NP or pronoun, thus keeping the original antecedent in the reader's short term or working memory.

#### **5.9.1.1 Reading and memory**

Siu (1988) and Chen (1990) found that readers remember hierarchically more important propositions in a text, which are likely to be topicalised or centralised grammatically or discursively, as is the case with the persistent topics in the texts in this study. Stevenson's (1993) claim that the goal of memorisation makes a reader focus on syntax, but the goal of comprehension makes the reader focus on propositions would indicate that readers of Chinese, as in the case of this study, would focus on discursal features such as topic persistence and continuity. As described in Section 2.1, Chinese does not have rich inflectional morphology, so remembering a Chinese text does not involve nearly as much focus on syntax anyway. All the memory energy can go into the propositions. Sachs (1967), comparing subjects' memory of active and passive forms in English, found that subjects were likely to be remembering a proposition rather than surface syntax. This could be the case in Chinese, especially with verbs which have flexible transitivity.

Zhang and Simon's (1985) finding that a normal Chinese short term memory span is up to seven non-homophonic characters (Section 2.2) may correspond roughly to a Chinese proposition. Topic hierarchy may also relate to what is retained in the memory and what is rejected. According to van Dijk and Kintsch (1983), short term memory is emptied at the end of each processing cycle, but important propositions are retained and these are likely to be 'high level' that is, those which relate to continuous, accessible, persistent topics.

#### **5.9.1.2 Reading and saccades**

As far as I know, there has been no experimental study of distribution of zero pronouns in relation to eye movements, but the texts in this study give the impression that intervals between the use of explicit NP or pronoun are similar in length to saccades. Sun and Feng (1999) maintain that eye movement patterns are controlled by linguistic information. Yang and McConkie (1999) found that saccade length in Chinese reading ranged from 1-14 units (0.5 - 7 characters) a length which may be similar to a Chinese proposition, or clause, and again they found that reading was controlled by linguistic information rather than visuo-spatial patterns of orthography. This is supported by a study by Inhoff, Liu and Tang (1999), who found that perceptual span in Chinese readers included the fixated character, plus up to two characters to its left and up to three characters to its right, a total of six.

#### **5.9.1.3 Reading and idea units**

The limitation on antecedent distance for zero pronouns may also be linked to some kind of cognitive 'unit' such as the 'idea unit' suggested by Chafe (1980, reviewed in Section 1.3). Chafe claims that eye movement provides evidence for the mechanism of focus of consciousness, following paths determined by the individual reader's schema. Chafe's 'idea unit' is a 'spurt of language' which consists, syntactically, of 'one verb with..... accompanying noun

phrases' (Chafe 1980). A series of several of these idea units is a centre of interest. In structural terms this discoursal description equates to Shi's description of the topic-comment structure as the basic syntactic structure in Chinese. There is some similarity, too, between Chafe's description and Fox's rhetorical structures, which are based on a hierarchy of propositions (Fox 1987).

The patterns of the topic chains found in the texts in this study suggest that it is also possible to equate the idea unit, or proposition not only to eye movements but also to short term memory capacity, and to antecedent distance for attenuated forms, such as zero pronouns.

#### **5.9.1.4 Reading and punctuation breaks**

As noted in Section 5.8 there may be some interaction between the distribution of zero pronouns, pronouns and punctuation. Chen (1986) found that zero pronouns often occurred in the same sentence as their antecedents, while pronouns were often separated from antecedents by sentence boundaries. As pointed out in Section 2.5.7.4, true spacing in Chinese text occurs only where there are punctuation marks, each of which takes up as much space as a single character. Chinese readers may not be guided by words, but by other multi-character units separated by punctuation marks. Where there is a physical break in the text there may also be a break in idea units, such as a non-conjoinable clause, a discontinuation of topic or theme, or a noteworthy new event or concept. Chinese readers have been found to fixate spaces, and it is likely that they are looking for syntactic or discoursal signals provided by the space, especially if the signal is reinforced by the use of an explicit anaphor. In many cases a zero anaphor may be used, however, and then the reader must rely entirely on the punctuation. It is also noticeable that low density, high frequency function words are 'skipped' (Wong 1995, Pollard 1995). This may not simply be negligence, but perception of the low



density characters between high density characters (Chen 199, see Section 2.2.4) as syntactic signals, taken in during a parfoveal preview (Inhoff, Liu and Tang 1999). Spacing in text cannot be dismissed as 'visuo-spatial layout' - linguistic structure is geared to perception and cognition through text layout, especially spaces.

#### **5.9.1.5 Reading and mental spaces**

The zero pronoun may also have some bearing on Fauconnier's (1994) theory of mental spaces, reviewed in Section 1.4. These are mental models of discourse guided, but not defined, by grammar (Fauconnier 1994). Mental spaces are 'called up' by grammatical signals, such as tense, mood, adverbs and so on, but are underspecified, not having a one to one correspondence with linguistic forms. In Chapter 2, the great economy of Chinese was discussed, showing that often Chinese prefers elliptical forms - not only the zero pronoun which is the focus of this study, but also clauses without connectors and even without verbs (Section 4.2.1.2).

I suggest that once a mental space has been accessed, in Fauconnier's terms, by explicit referents and by explicit adverbs, zero pronouns and other elliptical forms in Chinese in effect tell the reader or listener to stay in the space. When explicit referents are reintroduced in the same persistent topic chain these may be an aid to memory, reminding the reader which space it was, or may be signalling a change of event within the same space. When full NP referents are used in a topic chain without zero anaphors, this may signal the reader to access a new space, expand the space, or allow for the introduction of a new element into the space. This theory is not as neat and tidy as Chafe's idea units, but it does have some plausibility and deserves further investigation with reference to Chinese.



### **5.10 Zero pronouns, memory, saccades, punctuation, idea units, mental spaces and the reader**

Thus, the zero pronoun in Chinese is not a sign of linguistic impoverishment, nor is it a sign of direct ideation (Shen 1988). It is not a great mystery. As can be seen from the analysis of the Chinese source texts in the study, the zero pronoun plays a role in very clear, regular patterning of syntactic and discoursal structures of anaphoric choice by the writer, which send clear signals about anaphoric resolution to the reader. If patterns such as these can be seen in a text, then they can be taught to students. The analysis of subjects' parallel translations of the texts in this study shows that the majority of readers follow the signals. Where a minority of subjects are disrupted in their reading, the reasons, made clear in the translations, are to do with factors such as inappropriate schemata and inability to cope with abstraction.

## **Chapter 6: Implications for further research**

### **6.1 Revisiting the aims of the thesis**

#### **6.1.1 Primary aim**

The primary aim of this thesis as stated in Section 0.1.1 was to investigate some of the processes of reading Chinese text by analysing parallel translations of Chinese source texts into English target texts, focusing on the zero pronoun. The results of the study outlined in Chapter Five have made clear to what extent the interpretation of zero pronouns, and thereby the propositional representation of a text, can vary across a group of readers. The study has shown how certain features of text such as topic persistence, antecedent distance, the meaning of the verb, abstraction, multiplicity of arguments, clause boundaries and punctuation may affect readers' anaphoric resolution. It has also shown how reader characteristics such as background knowledge and personal schema may affect reading.

What the study has not covered is how these factors as revealed in translation correlate with tested reading ability and tested language ability. It is likely that achievement of optimal solutions in anaphoric resolution would correlate with language ability, reading ability and probably social factors such as age and experience.

#### **6.1.2 Secondary aim**

The secondary aim of the thesis was to show how translation from L2 to L1 can give evidence of reading comprehension, and therefore, if sensitively applied could be a useful testing tool. The subjects in this study were bilingual, that is conversant with both Chinese and English, but covered a wide range of ability. For the majority of them English is the language of habitual use.

The study shows that translation of a whole text reveals aspects of reading comprehension not necessarily covered by short answer questions, in this case the understanding of reference patterns throughout the text. As explained in Chapters Two, Three and Four, this particular problem is of crucial importance in the case of Chinese which does not have rich morphological inflection. Translation between languages as dissimilar as Chinese and English provides rich evidence of cognitive processes, from orthographical miscues at the lowest level, syntactic and pragmatic miscues and misapplication of schemata at the highest levels of reading. Some of the problems seen in translations from L2 to L1 may be due to a lack of awareness of typological differences between the two languages. What this study does not show is whether any degree of training in translation, even as a minor exercise, contributes positively to a reader/translator's ability to comprehend a text. Yan (2000) suggests that it does, and this requires further statistical investigation.

## **6.2 The extent of the role of translation in analysing reading processes**

In Chapters One and Two a range of theoretical and experimental studies of English reading and Chinese reading were reviewed. The review of the theoretical studies illustrated the range of approaches to reading and the teaching of reading. This study confirms and corroborates the view that reading is an interactive process in which information at every level triggers and confirms information at every other level. The experimental studies reviewed included eye movement studies, recall studies, cloze tests, and violation experiments. The results of this study corroborate some of these experiments, showing that there may be some mileage in using translation as a complementary means of studying reading processes.

What remains to be evaluated is the degree to which translation confirms experimental findings, and what particular aspects of reading are revealed in

translation of continuous text but not revealed in reductionist experiments. This study shows that one of the main advantages of translation from L2 to L1 is that it allows an unlimited range of response which is still quantifiable and analysable, whereas on-line experiments may only offer a limited range of responses.

### **6.3 What translation reveals about comprehending reference patterns in a text**

Taking the analysis of optimal, proximal and non-solutions of anaphoric resolution in this study together with the analysis of subjects' continuity of reference, it is evident that translation reveals a great deal about understanding of reference chains in a text. What appear to be quite accessible, unambiguous anaphors are not always understood. The patterns of reference in the texts show distinct features such as topic persistence, antecedent distance, backgrounding, and so on, which are confirmed by previous research, notably that of Li and Thompson (1979), Givon (1983), Chen (1986) and Tao (1993). These patterns are teachable, and are probably best taught by using selected texts of different genres, showing students the outstanding patterns of Chinese text such as the attenuation associated with the persistent central topic, the full NP associated with backgrounded theme or non-central protagonists, and the use of less attenuated forms, ie pronouns, to assist short term memory in long topic chains, or to indicate highlighting, as in the case of new events or episodes. Comparison needs to be made of teachers' and students' reactions to translation as a reading comprehension exercise as compared with short answer questions, and the relative effectiveness of the two methods.

As pointed out previously, the teaching of Chinese characters is not covered in this thesis, which assumes that perception of print is automatised by the stage, or age, of the subjects in this study. However, low level miscues are

revealed in translation to L1 and could be dealt with as part of an interactive approach.

Models such as that of van Dijk and Kintsch (1978), which shows how text structure is built up from linking of propositions, or that of Chafe (1980), which describes the linking of idea units into a centre of interest carried through given and new information provide frameworks for the successful teaching of reading. Discourse based approaches may be particularly suited to Chinese with its lack of strictly grammatical information, and have been successfully developed and adapted to that end by Li and Thompson (1979), Xu (1990), Huang (1994), Chen (1986) and Tao (1993). These discourse based approaches are more useful pedagogically than generativist approaches, which deal only with intra-sentential reference. Moreover the discourse analysis approach is better suited to Chinese, in which the recoverability of zero pronouns is not aided by a rich inflectional morphology.

One crucial aspect of reading Chinese text is recognising how the visuo-spatial features, such as density and spacing, can help. Assuming that the reading of Chinese characters is largely automatised, the important thing for the reader to notice in a text is the spacing. The spaces in Chinese text are associated with punctuation breaks, and therefore give some indication of clause structure. Good readers of Chinese use the parafoveal information of a space and its punctuation mark to predict, for example, contrast, unexpectedness, new event, or new protagonist. The pattern of reference is closely linked to the pattern of clause structure.

While it appears that in Chinese reading lexical access precedes any other form of access (Chen 1999), awareness of syntactic structures is also vital to understanding. The topic structure in Chinese, which includes anaphoric devices, can arguably be described as both a discoursal structure (Li and

Thompson 1979) and a syntactic structure (Shi 1989). While syntactic analysis helps, pragmatic knowledge is sometimes the only way to disambiguate a zero pronoun. The role of the predicate, or verb, is vital in these cases, and students can be taught to be aware of verb meaning as an aid to anaphoric resolution.

### **6.3.1 Using translation as a reading comprehension exercise**

The degree of usefulness will depend firstly on the nature of the text. Tricky, obscure texts will teach nothing except negative attitudes. It is important that texts for translation are selected in line with the students' current stage of development and as exemplars of structures or vocabulary on the syllabus currently in focus, new items being clear from context.

'Reading for thought' (Huey 1908), reading for meaning, '*la saisie du sens hors langue*' (Lederer 1976) are core concepts for reading and for translation. The notion of equivalence (Section 3.2.2) and its companion '*sourcier*' or literal translation (Section 3.2.2) are not effective means of gaining understanding of a text. Students need to be prepared for new vocabulary, provided with a suitable schema by means of a title, and encouraged to read deeply and globally.

### **6.3.2 Parallel translations**

Analysis of student translators' parallel translations (Toury 1995), literary parallel translations (Hofstadter 1987) and translation corpora (Baker 2001) show how varied, yet how quantifiable different readers' approaches can be, and reveal much about translators' cognitive processes. Statistical analysis of a large number of translations, as carried out in this study, shows how the majority of subjects will achieve a norm, or optimal solution, while the proximal and non-solutions of the minority reveal much about how mistakes are made. If this is true at a professional level, it should surely filter down to

the lower levels of language education from which succeeding generations of professional translators emerge.

Chinese is widely perceived as mysterious and untranslatable - not a helpful attitude for learners of the language. The analysis of translation, as carried out in this study, by revealing the patterns within Chinese texts, and by showing that there are norms, or at least optimal solutions in comprehension and translation, may help to dispel the myths.

Explicitation ( adding elements in L1 which are not apparent in the L2 source text) is a case in point. Whereas Blum-Kulka (1986) believes this to be inherent in translation, my analysis shows that it is not necessarily. According to my data, intuitive translators appear to use explicitation either when necessary grammatically or when setting out clearly their perception of the propositional representation of the text. The use of translation corpora could assist enormously by quantifying the use of obligatory explicitation and stylistic explicitation, and might establish to what extent it is inherent behaviour, conscious behaviour, or simply language specific.

### **6.3.3 Translation in language testing**

As long as students have been carefully trained to translate appropriate, meaningful texts which contain known, accessible and useful structures and vocabulary, there is no reason why a similar approach should not be used for language testing. As pointed out previously, it may be useful diagnostically, and may reveal errors which short answer questions or multiple choice questions might not necessarily reveal. Some language educators are opposed to translation as a test type. However, objections to translation seem to be based mainly on a perception of it as it was several decades ago - random, tricky, literary, irrelevant texts designed to catch you out. It is neither necessary nor desirable to present students with too difficult texts.



Apart from the paper by Källkvist comparing essays and translations (Källkvist 1998), there appears to be no research which objectively evaluates translation as a test type. Much could be done to clarify the situation if objective statistical analysis of comparable subjects taking conventional target language reading comprehension tests and L2 to L1 translation tests were carried out. This could be done on the basis of the type of analysis carried out in this study and the trial vocabulary analysis described in Sections 3.6.3.4.1 and 3.6.3.4.2.

As pointed out in Section 3.6.3.4, the main difficulty of translation as a test type is the problem of objective, standardised marking. It is customary to mark essays according to scalar systems, and this could also be applied to translation. The findings of this study, as set out in Chapter Five, suggest that reader/translators' solutions fall into 4 bands - optimal solution, proximal solution, non-solution and omission. This might vary according to the languages involved, the type of text and the level of the students. From these bands effective objective descriptors could be formulated and markers could be trained to apply criteria as they are currently trained for conventional test types, including essay writing. Year on year comparison with other test types would provide feedback on reliability and validity.

#### **6.4 The need for joined-up research into features of text and reader behaviour**

The findings of this study as set out in Chapter Five suggest that there may be close, specific relationships between length of saccade as described in eye movement studies, short term or working memory capacity, idea units, the length of clauses, the occurrence of punctuation breaks and antecedent distance. When a topic persists over several clauses, even if central and thus referred to by means of zero anaphora, overt anaphors are used at intervals to remind the reader of the topic. Antecedent distance - the distance between a

zero anaphor and its NP or pronoun antecedent - is never very far, and the texts used in this study suggest that antecedent distance equates with the amount of material a reader's short term memory can cope with. This in turn appears to be about the same as an average saccade (between 2 and 14 Chinese characters, but average 7). A saccade of 7 characters would be roughly equivalent in length to a proposition or a simple clause. This would not be difficult to verify experimentally.

A comprehensive linking of the factors described above would provide definitive information about the relationship between timing, tracking and syntactic analysis. Now that PET is available, such research could be enhanced by discovering what neurological connections are made as these reading processes take place. If this type of experimental research were carried out on both reading and translation, this might reveal what concepts are being processed, and where, how quickly, and by means of what syntactic analysis they are processed, thus closing some of the gaps, if not all, in the theory and practice of reading comprehension.

### **6.5 A summary of possible ways forward**

1. Studying the relationship between ability in L2, reading ability in L2, as tested by conventional means, and reading ability in L2 as tested by translation into L1.
2. Investigating whether or not some degree of training in translation improves reading ability in L2, and whether cross-linguistic awareness can be developed by means of translation exercises.
3. Investigating the extent to which studies of reading by means of translation corroborate or refute experimental studies of reading and what particular gaps in the research translation might fill.

4. Classroom research into diagnosis of reading problems by means of translation and development of remedial work on those problems carried out by study of appropriate texts.
5. In the case of the zero anaphor, a statistical analysis of intuitive reader/translators' use of explicitation in translation from L2 to L1, to clarify the extent to which explicitation is obligatory, stylistic, or evidence of cognitive processes.
6. Objective evaluation of translation as a test type compared with conventional L2 reading comprehension tests.
7. In the particular case of Chinese reading, co-ordination of research into saccade length and short term or working memory capacity in the reader, in relation to idea units, clause length, punctuation breaks and antecedent distance in the text.

This study suggests firstly, that the processes of reading comprehension can be revealed by analysing parallel translations of L2 to L1, as shown here by subjects' interpretation of the Chinese zero anaphor. Secondly, the study suggests that translation from L2 to L1 provides evidence of comprehension, and may therefore be useful as a test type. Further research as outlined above would give more solid evidence for or against these suggestions.

## References

- Adams, M.J. (1990) *Beginning to read: the new phonics in context*. Oxford. Heinemann.
- Alderson, J.C and Urquhart, A.H. (eds.) (1984) *Reading in a foreign language*. London. Longman.
- Aoun, J. (1985) *A grammar of anaphora*. Cambridge MA. The MIT Press.
- Au, T. K-F. (1992) Cross-linguistic research on language and cognition: methodological challenges. In Chen and Tzeng (eds.) (1992) 367-381.
- Bachmann, L.F. (1990) *Fundamental considerations in language testing*. Oxford. OUP.
- Baddely, A. and Lewis, V. (1991) Inner active processes in reading: the inner voice, the inner ear and the inner eye. In Lesgold and Perfetti (eds.) (1981) 107-129.
- Baddeley, A. D., Logie, R.H. and Ellis N.C. (1988) Characteristics of developmental dyslexia. *Cognition* 29. 197-228.
- Baker, M. (ed.) (1998) *Routledge Encyclopedia of translation studies*. London. Routledge.
- Baker, M. (2001) Translation Corpora. Paper given to Edinburgh Institute for the Advanced Study of Islam and the Middle East, University of Edinburgh 23.2.01
- Balota, D., Flores d'Arcais, G. and Rayner, K. (eds.) (1990) *Comprehension processes in reading*. Hillsdale, NJ. Erlbaum.
- Bartlett, F.C. (1932) *Remembering: a study in experimental and social psychology*. London. Cambridge University Press.
- Bassnett, S. and Lefevere, A. (eds.) (1990) *Translation, history and culture*. London. Pinter Publishers Ltd.
- Berman, R.A. (1984) Syntactic components of the foreign language reading process. In Alderson and Urquhart (eds.) (1984). 139-156.
- Blakemore D. (2000) The pragmatics of and-conjunctions: the non-narrative cases. Paper given to University of Edinburgh Linguistics Circle.

- Bloom, A. (1981) *The linguistic shaping of thought: a study in the impact of language on thinking in China and the West*. Hillsdale, NJ. Lawrence Erlbaum Associates.
- Bond, M.H. (ed.) (1996) *The handbook of Chinese psychology*. Hong Kong. OUP.
- Brown, G. and Yule, G. (1983) *Discourse analysis*. Cambridge. CUP.
- Brown, K. and Miller, J. (eds.) (1999) *Concise Encyclopaedia of grammatical categories*. Oxford. Pergamon.
- Bryant, P (1995) Phonological and grammatical skills in learning to read. In de Gelder and Morais (eds.) (1995) 249-266.
- Burmeister, H. and Rounds, P. (eds.) (1990) *Variability in second language acquisition*. Proceedings of the 10th meeting of the Second Language Research Forum. Eugene. University of Oregon.
- Candlin, C.N. (1986) Explaining communicative competence limits of testability? In Stansfield (ed.) (1986).
- Carrell, P.L. (1988) Interactive text processing: implications for ESL/second language reading classrooms. In Carrell, Devine and Eskey (eds.) (1988) 101-113.
- Carrell, P.L. Devine, J. and Eskey, D.E. (eds.) (1988) *Interactive approaches to second language reading*. Cambridge. CUP.
- Catford, J.C. (1965) *A linguistic theory of translation*. London. OUP.
- Cayley, J. (1995) Chinese classical poetry. In Chan and Pollard (eds.) (1995). 759-772.
- Chafe, W.L. (1976) Givenness, contrastiveness, definiteness, subject, topic and point of view. In Li (ed.) (1976) 25-55.
- Chafe, W.L. (1980) *The Pear Stories: cognitive, cultural and linguistic aspects of narrative production*. Norwood, NJ. Ablex Publishing Corporation.
- Chan, S.-Y., and Pollard, D. E. (eds.) (1995) *An encyclopaedia of translation. Chinese-English, English-Chinese*. Hong Kong. The Chinese University Press.
- Chao, Yuen Ren. (1968) *A grammar of spoken Chinese*. Berkeley. University of California Press.

- Chau, S. S-C. (1984) Aspects of translation pedagogy: the grammatical, cultural and interpretive teaching models. Unpublished Ph.D. Thesis. University of Edinburgh.
- Chen, C.M. (1981) Perception of Chinese characters. *Acta Psychologica Taiwanica*. 23. 137-358.
- Chen, D-X. (1990) Reading in Chinese and English: effects of formal schema and metacognition. In Burmeister and Rounds (eds.) (1990).
- Chen, H-C. (1992) Reading comprehension in Chinese: implications from character reading times. In Chen and Tzeng (eds.) (1992) 175-205.
- Chen, H-C. (1996) Chinese reading and comprehension: a cognitive psychology perspective. In Bond (ed.) (1996) 43-62.
- Chen, H-C. (ed.) (1997) *Cognitive processing of Chinese and related Asian Languages*. Hong Kong. The Chinese University Press.
- Chen, H-C. (1999) How do readers of Chinese process words during reading for comprehension? In Wang, Inhoff and Chen (1999) 257-278.
- Chen, H-C. and Tzeng, O.J.L. (eds.) (1992) *Language processing in Chinese*. Amsterdam. North-Holland (Elsevier Science Publishers. BV.).
- Chen, H-C. and Zhou, X-L. (1999) Processing East Asian languages: an introduction. *Language and cognitive processes*. 1999.14.
- Chen, J-Y. (1999) Word recognition during the reading of Chinese sentences: evidence from studying the word superiority effect. In Wang, Inhoff and Chen (eds.) (1999). 239-256.
- Chen, M.J., Yung, Y.F., and Ng, T.W. (1988) The effect of context on perception of Chinese characters. In Liu, Chen and Chen (eds.) (1988) 27-39.
- Chen, P. (1986) Referent introducing and tracking in Chinese narrative. Ph.D. dissertation. University of California at Los Angeles.
- Chen, P. (1987) Hanyu lingxing huizhi de huayu fenxi. *Zhongguo yuwen*. 1987 No. 5.
- Chen, S., Tzeng, O.J.L. and Bates, E. (1990) Sentence interpretation in Chinese aphasia. In Burmeister and Rounds (eds.) (1990).

- Chomsky, N. (1982) *Lectures on government and binding. The Pisa Lectures.* Dordrecht. Foris Publications.
- Clancy, Patricia M. (1980) Referential choice in English and Japanese narrative in discourse. In Chafe (ed.) (1980) 127-202.
- Clark, H.H. (1978) Inferring what is meant. In Levelt and Flores d'Arcais (eds.) (1978) 259-322.
- Cohen, A. Glasman, H., Rosenbaum-Cohen P.R. Ferrara, J. and Fine, J. (1988) Reading English for specialised purposes: discourse analysis and the use of student informants. In Carrell, Devine and Eskey (eds.) (1988) 152-167.
- Coltheart, M., Patterson, K. And Marshall, J.C. (1980) *Deep dyslexia.* London. Routledge and Kegan Paul.
- Comrie, B. (1981) *Language universals and linguistic typology: syntax and morphology.* Oxford. Blackwell.
- Connor, U. (1984) Recall of text: differences between first and second language readers. *TESOL Quarterly* 18 (2). 239-256.
- Connor, U. (1996) *Contrastive rhetoric.* Cambridge. CUP.
- Cui, Y. and Chen Y.M. (1997) Predictive inference in reading. In Chen, H-C. (ed.) (1997) 287-297.
- Danchev, A. (1983) The controversy over translation in foreign language teaching. In *Translation in Foreign Language Teaching Round Table.* Paris. FIT-UNESCO.
- de Francis, J. (1989) *Visible speech: the diverse oneness of writing systems.* Honolulu. University of Hawaii.
- de Gelder, B. and Morais, J. (eds.) (1995) *Speech and reading: a comparative approach.* Hove. Erlbaum (UK) Taylor and Francis.
- Downing, P. and Noonan, M. (eds.) (1995) *Word order in discourse.* Amsterdam. John Benjamins Publishing Company.
- Duff, A. (1981) *The third language: recurrent problems of translation into English.* Oxford. Pergamon.



Edexcel. (1998, 2000) *London Examinations Advanced Level GCE Chinese*. London. Edexcel.

Edexcel. (2000) *University of London GCE Examinations Oriental and African Languages Syllabus*. London. Edexcel.

Faerch, C. and Kasper, G. (1986) 'One learner – two languages:' investigating types of interlanguage knowledge. In House and Blum-Kulka (eds.) (1986).

Faerch C. And Kaspar, G. (eds.) (1987) *Introspections in second language learning research*. Clevedon. Multilingual Matters.

Fauconnier G. (1994) *Mental spaces: aspects of meaning construction in natural language*. Cambridge. CUP.

Fillmore, C. J. (ed.) (1971) *Studies in linguistic semantics*. New York. Holt, Rinehart and Winston.

Fillmore, C. J. (1971) Verbs of judging. In Fillmore, C.J. (1971) 273-296.

Fillmore, C.J. (1982) Frame Semantics. In Linguistic Society of Korea (eds.) (1982).

Fillmore, C.J. (1982) Monitoring the reading process. In Linguistic Society of Korea (eds.) (1982).

Fox, B. A. (1987) *Discourse structure and anaphora: written and conversational English*. Cambridge. CUP.

Fox, B.A. (ed.) (1996) *Studies in anaphora*. Typological studies in language, 33. Amsterdam. John Benjamins Publishing Company.

Garrod, S. (1986) Language comprehension in context: a psychological perspective. *Applied Linguistics* 3. 226-238.

Garrod, S. and Sanford A.J. (1977) Interpreting anaphoric relations: the integration of semantic information while reading. *Journal of Verbal Learning and Verbal Behaviour* 16. 77-90.

Gatenby, E.V. (1967) Translation in the classroom. In Lee, W.R. (ed.) (1967) *ELT selections 2: articles from the journal 'English Language Teaching.'* London. OUP.

- Gerver, D. and Wallace Sinaiko (eds.) (1977) *Language interpretation and communication*. New York. Plenum Press. NATO Conference Series.
- Givon, T. (ed.) (1979) *Syntax and Semantics Vol. 12: Discourse and syntax*. London. Academic Press.
- Givon, T. (ed.) (1983) *Topic continuity and discourse*. Amsterdam. John Benjamins Publishing Company.
- Glushko, R. J. (1979) The organization and activation of orthographic knowledge in reading aloud. *Journal of experimental psychology: human perception and performance*. 5. 674-691.
- Goodman, KS. (1967) Reading: a psycholinguistic guessing game. *Journal of the Reading Specialist* 6. 126-135.
- Grabe, W. (1988) Re-assessing the term interactive. In Carrell, Devine and Eskey (eds.) (1988) 56-70.
- Harris, B. (1977) The importance of natural translation. *Working Papers on Bilingualism*, 12. 96-114.
- Harris, B. (1992) Natural translation: a reply to Hans Krings. In *Target* 1992, Vol. 4, Pt. 1. 97-103.
- Harris, B. and Sherwood, B. (1977) Translating as an innate skill. In Gerver, D., and Wallace Sinaiko (eds.) (1977) 155-170.
- Hartmann, R.K. (1995) Contrastive textology, bilingual lexicography and translation. In Chan and Pollard (1995) 505-518.
- Hermon G. and Yoon, J. (1989) The licensing and identification of pro and the typology of Agr. *Papers from the regional meeting of the Chicago Linguistic Society* 25. 174-192.
- Hofstadter, D.R. (1997) *Le ton beau de Marot: in praise of the music of language*. New York . Basic Books.
- Hoosain, R. (1991) *Psycholinguistic implications for linguistic relativity: a case study of Chinese*. Hillsdale, NJ. Lawrence Erlbaum Associates.
- Hoosain, R. (1995) Getting at the meaning of logographic and alphabetic scripts. In Taylor and Olson (eds.) (1995) 131-145.

- House, J. and Blum-Kulka, S. (eds.) (1986) *Interlingual and intercultural communication: discourse and cognition in translation and second language acquisition studies*. Tübingen. Gunter Narr Verlag.
- Howatt, A.P.R. (1984) *A history of English language teaching*. Oxford. OUP.
- Hu, J-H. (1999) The finite vs. nonfinite distinction and null subjects in Chinese. Paper given at Linguistics Society of Hong Kong Annual Research Forum.
- Huang, C-T.J. (1989) Pro-drop in Chinese: a generalized control theory. In Jaeggli, O.A. and Safir, K.J. (1989). 185-214.
- Huang, H.F. (1978) *Zhong-ying fanyi: lilun yu shijian*. Taibei. Wenhong Chuban Youxian Gongsu.
- Huang, Y. (1994) *The syntax and pragmatics of anaphora*. Cambridge. CUP.
- Huang, Y. (2000) *Anaphora: a cross-linguistic study*. Oxford. OUP.
- Huang, Y-S. (1995) Form and spirit. In Chan and Pollard (eds.) (1995) 277-287.
- Hudson T. (1982) The effects of induced schemata on the 'short circuit' in L2 reading: non-decoding factors in L2 reading performance. *Language Learning* 32 (1): 1-31.
- Huey, E.B. (1908) (1968)) *The psychology and pedagogy of reading*. Cambridge MA. MIT Press.
- Inhoff, A. and Liu, W. M. (1997) The range of effective vision during the reading of sentences. In Chen (ed.) (1997) 243-265.
- Inhoff, A.W., Liu, W.M. AND Tang, Z-H. (1999) Use of prelexical and lexical information during Chinese sentence reading: evidence from eye movement studies. In Wang, Inhoff and Chen (eds.) (1999). 223-238.
- Jaeggli, O.A. and Safir, K.J. (1989) *The null subject parameter*. Dordrecht. Kluwer.
- Just, M.A. and Carpenter, P.A. (1976) Eye fixations and cognitive processes. *Cognitive psychology*. 1976. 8. 441-480.
- Just M.A. and Carpenter, P.A. (1992) A capacity theory of comprehension: individual differences in working memory. *Psychological Review* 99. 122-149.

Just, MA., Carpenter, P.A. and Woolley (1982) Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General* III. 228-238.

Källkvist, M. (1998) How different are the results of translation tasks? A study of lexical errors. In Malmkjaer K. (1998) 77-87.

Kang B-W. (1999) Tongzuzi jiaoxue baogao – jianping xiaoxue zhongwenke changyong zibiao. International Language in Education Conference 1999. Chinese University of Hong Kong.

Kaplan, R.B. (1966) Cultural thought patterns in intercultural education. *Language Learning* 16. 1-20.

Kintsch, W. and van Dijk, T.A. (1978) Toward a model of text comprehension and production. *Psychological Review* 85. 363-394.

Klaudy, K. (1998) Explication. In Baker (ed.) (1998) 80-84.

Klein-Braley, C. and Franklin, P. (1998) 'The foreigner in the refrigerator:' remarks about teaching translation to university students of foreign languages. In Malmkjaer (ed.) (1998) 53-61.

Krings, H.P. (1987) The use of introspective data in translation. In Faerch and Kasper (ed.) (1987).

Kuno, S. and Kaburaki, E. (1977) Empathy and syntax. *Linguistic Inquiry* 8. 627-672.

Laberge, D. and Samuels, J. (eds.) (1977) *Basic processes in reading: perception and comprehension*. Hillsdale, NJ. Erlbaum.

Lam, J. K-M. (1995) Thinking aloud protocol. In Chan and Pollard (eds.) (1995) 904-917.

La Polla R.J. (1995) Pragmatic relations and word order in Chinese. In Downing and Noonan (eds.) (1995) 297-329.

Lederer, S. (1973) Transcoder ou réexprimer? In Seleskovitch, D. and Lederer, S. (1993) 15-36.

Lederer, S. (1976) Implicite et explicite. In Seleskovitch and Lederer, (1993) 37-71.

- Lee, T.H-T. (1992) (ed.) *Research on Chinese Linguistics in Hong Kong*. Hong Kong. Linguistic Society of Hong Kong.
- Lee, W.L., Wee, G.C. and Wong, E. (1987) Cognitive correlates of reading in English and Chinese among bilingual children in Singapore. Paper presented at the Annual Meeting of the Orton Society of Dyslexic Studies. San Francisco.
- Lefevere A. (1995) Factors of poetic translation. In Chan and Pollard (eds.) (1995) 747-757.
- Leong C.K. and Hsia, S. (1996) Cross-linguistic constraints on Chinese students learning English. In Bond (ed.) (1996) 63-78.
- Lesgold A.M. and Perfetti C.A. (1981) *Interactive processes in reading*. Hillsdale, NJ. Lawrence Erlbaum Associates.
- Levelt, W.J.M. and Flores d'Arcais G.B. (eds.) (1978) *Studies in the perception of language*. Chichester. John Wiley and Sons.
- Levinson, S.C. (1987) Pragmatics and the grammar of anaphora: a partial pragmatic reduction of binding and control phenomena. *Journal of Linguistics* 23. 379-434.
- Li, C.N. (ed.) (1976) *Subject and topic*. London. Academic Press.
- Li, C. N. and Thompson, S.A. (1979) Third-person pronouns and zero-anaphora in Chinese discourse. In Givon (ed.) (1979) 311-335.
- Li, C.N. and Thompson, S.A. (1981) *Mandarin Chinese: a functional reference grammar*. Berkeley. University of California Press.
- Li, P., Bates, E., Liu, H. and MacWhinney B. (1992) Cues as functional constraints on sentence processing in Chinese. In Chen and Tzeng (eds.) (1992) 207-234.
- Lin, A.M-Y. and Akamotsu, N. (1997) The learnability and psychological processing of reading in Chinese and reading in English. In Chen (ed.) (1997) 369-388.
- Lin Yutang. (1961) *The importance of understanding*. London. Heinemann.
- Linguistic Society of Korea (eds.) SICOL-1981 (1982) *Linguistics in the morning calm*. Seoul. Hanshin Publishing Company.

- Liu, I.-M. (1995) Script factors that affect literacy: alphabetic vs. logographic languages. In Taylor and Olson (eds.) (1995) 145-160.
- Liu, I.M., Chen, H.C. Chen, M.J. (1988) *Cognitive aspects of the Chinese language*. Hong Kong. Asian Research Service.
- Liu, I.M., Yeh, J.S., Wang, L.H., and Chang, Y.K. (1974) Effects of arranging Chinese words as units on reading efficiency. *Acta Psychologica Taiwanica* 16. 25-32.
- Liu, Q.M. (1995) Grammar and translation. In Chan and Pollard (ed.) (1995) 301-316.
- Lu, D. (2000) Message transfer between L1 and L2: reflections on bilingual competence. Paper given at International Language in Education Conference (ILEC) 2000, Hong Kong.
- Lu, S.-X. (1986) Hanyu jufa de linghuoxing. *Zhongguo Yuwen* 1986 Vol.1.
- Lust, B. (1981) Constraints on anaphora in child language: a predication for a universal. In Takavolian (ed.) (1981).
- Lust, B., Lovelend, K. and Kornet, R. (1980) The development of anaphora in first language: syntactic and pragmatic constraints. *Linguistic Analysis* 6.4. 359-391.
- Malmkjaer, K. (ed.) (1998) *Translation and language teaching: language teaching and translation*. Manchester. St. Jerome Publishing.
- Marot, Clément. (1990-1993) *Oeuvres poétiques complètes*. Paris. Bordas.
- Mattingly I.G. and Hsiao, P.-L. (1997) Constituent superiority in Chinese. In Chen H.-C. (1997) 207-218.
- Miao, Q.Z. (1986) Xiandai hanyu bianzhangzhong zhitong de biaoda. *Zhongguo Yuwen* 1986.2.
- Miao, X.C. (1981) Word order and semantic strategies in Chinese sentence comprehension. *International Journal of Psycholinguistics* 8. 109-122.
- Miao, X.C. (1996a) The effect of sentence semantics and distance between pronoun and its antecedent on pronoun processing. *Psychological Science* 2. 71-74.

Miao X.C. (1996b) A study of semantic and grammatical factors influencing pronoun processing. *Acta Psychologica Sinica* 4. 352-358.

Miao, X.C. (1999) Sentence understanding in Chinese. In Wang, Inhoff and Chen (eds.) (1999) 279-296.

Miao, X.C. and Song, Z.G. (1995) The influence of verb meaning and sentence grammar on pronoun processing. *Psychological Science* 4. 197-200.

Miller, G.A. (1956) The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review* 63. 81-87.

Minsky, M. (1975) A framework for representing knowledge. In Winston, (ed.) (1995) 211-277.

Muhawi, I. (1998) On translating Mahmoud Darwish – some issues in translation theory and practice. Paper given at 'The Translator at work' conference, Edinburgh Institute for the Advanced Study of Islam and the Middle East.

Nabokov, V. (1955) Problems of translation: Onegin in English. *Partisan Review* 22. No.4.

Newmark, P. (1991) *About translation*. Clevedon. Multilingual Matters.

Newson, D. (1998) Translation in foreign language learning. In Malmkjaer (ed.) (1998) 63-68.

Nida, E. A. (1964) *Towards a science of translation: with special reference to principles and procedures involved in Bible translating*. Leiden. Brill.

Oller, J.W. Jr. (1979) *Language tests at school: a pragmatic approach*. London. Longman.

Peng, D-L. and Liu, S-L. (1993) Syntactic and semantic analysis in Chinese sentence comprehension. *Acta Psychologica Sinica* 2. 132-139.

Perfetti, C.A. (1985) *Reading ability*. New York. Longman.

Perfetti, C.A. (1990) The co-operative language process: semantic influences in an autonomous syntax. In Balota, Flores d'Arcais and Rayner (eds.) (1990).

Perfetti, C.A. and McCutcheon, D. (1987) Schooled language competence: linguistic abilities in reading and writing. In Rosenberg (ed.) (1987) 105-137.



Pollard, D.E. (1995) Empty words: modal adverbs. In Chan and Pollard (1995) 216-222.

Qian Gechuan (no date) *Fanyi de jiben zhishi (The ABC of translation)*. Hunan. Hunan Kexue Jishu Chubanshe.

Rayner, K. and Pollatsek, A. (1989) *The psychology of reading*. Hillsdale, NJ. Lawrence Erlbaum Associates.

Redvers-Jones, F. (1995) Translation in language teaching. In Chan and Pollard (eds.) (1995) 487-504.

Reichman R. (1981) *Plain speaking: a theory and grammar of spontaneous discourse*. Cambridge Ma. Bolt, Beranek and Newman Report 4681.

Reinhart, T. (1983) *Anaphora and semantic interpretation*. London. Croom Helm.

Rosenberg, D.S. (ed.) (1987) *Advances in Psycholinguistics. Vol. 2 Reading, writing and language learning*. Cambridge. CUP.

Rumelhart D. (1977) Understanding and summarizing brief stories. In Laberge and Samuels (eds.) (1977) 265-303.

Rumelhart, D.E. (1980) Schemata: the building blocks of cognition. In Spiro, Bruce and Brewer (eds.) (1980) 35-58.

Rutter, M. And Herson, L. (eds.) (1976) *Child psychiatry: modern approaches*. London. Blackwell Scientific.

Rutter, M. And Yule, W. (1976) Reading difficulties. In Rutter and Herson (eds.) (1976).

Sachs, J.S. (1967) Recognition memory for syntactic and semantic aspects of connected discourse. *Perception and psychophysics* 2. 437-442.

Samuels S.J. and Kamil M.L. (1988) Models of the reading process. In Carrell, Devine and Eskey (eds.) (1988) 22-36.

Schank, R. C. (ed.) (1975) *Conceptual information processing*. Amsterdam. Elsevier.

Schank R.C. and Abelson R.P. (1977) *Scripts, plans, goals and understanding: an inquiry into human knowledge structures*. Hillsdale, NJ. Lawrence Erlbaum.

- Schulte R. and Biguenet, J. (eds.) (1992) *Theories of translation: an anthology of essays from Dryden to Derrida*. Chicago. University of Chicago.
- Seleskovitch, D. and Lederer, M. (1993) *Interpreter pour traduire*. Paris. Didier Erudition.
- Seleskovitch, D. (1976) De l'expérience aux concepts. In Seleskovitch and Lederer (1993) 72-103.
- Sells P. and Wasow, T. (1999) Anaphora. In Brown and Miller (eds.) (1999) 18-25.
- Senders J.W., Fisher D.F. and Monty, R.A. (eds.) (1978) *Eye movements and higher psychological functions*. Hillsdale NJ. Lawrence Erlbaum.
- Shen, E. (1927) An analysis of eye movements in the reading of Chinese. *Journal of experimental psychology* 10. 158-183.
- Shen X-L. (1988) *Zhongguo yuyan de jiegou yu renwen jingshen*. Beijing. Guangming Ribao Chubanshe.
- Shi D-X. (1989) Topic chain as a syntactic category. *Journal of Chinese Linguistics* Vol. 17. 223-262.
- Sinn, E. (1995) Yan Fu. In Chan and Pollard (eds.) (1995). 429-447.
- Siu, P-K. (1988) Processing of thematic information in Chinese texts. In Liu, Chen and Chen (eds.) (1988). 163-170.
- Snell-Hornby, M. (1988) *Translation studies: an integrated approach*. Amsterdam. John Benjamins Publishing Company.
- Spiro, J.R., Bruce, B.C. and Brewer, W.E. (eds.) *Theoretical issues in reading comprehension*. Hillsdale, NJ. Erlbaum.
- Stanovich, K.E. (1980) Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly* 16. 32-71.
- Stanovich, K.E. (1981) Attentional and automatic context effects in reading. In Lesgold and Perfetti (eds.) (1981) 241-267.

- Stansfield C.W. (ed.) (1986) *Toward communicative competence testing: proceedings of the second TOEFL invitational conference*. Princeton NJ. Educational Testing Service.
- Stern, J.A. (1978) Eye movements, reading and cognition. In Senders, Fisher, and Monty (eds.) (1978).
- Stevenson H.W. (1984) Orthography and reading disabilities. *Journal of Learning Disabilities* 17. 296-301.
- Stevenson, H. and Lee, S-W. (1996) The academic achievement of Chinese students. In Bond (ed.) (1996). 124-142.
- Stevenson, H.W., Stigler, J.W. and Lee, S-Y. (1985) Cognitive performance and academic achievement of Japanese, Chinese and American children. *Child Development* 56. 718-734.
- Stevenson, H.W., Stigler, J.W., Lucker, W., Lee, S-Y. Hsu, C.C. and Kitamura, S. (1982) Reading Disabilities: the case of Chinese, Japanese and English. *Child Development* 53. 1164-1181.
- Stevenson, R.J. (1993) *Language, thought and representation*. Chichester. J. Wiley and Sons.
- Stibbard, R. (1998) The principled use of translation in foreign language teaching. In Malmkjaer (ed.) (1998) 69-76.
- Su, S.J., Soong, W.T. and Hsu, C.C. (1984) Dyslexia in Chinese children: the writing errors and WISC-R patterns. *Acta Psychologica Taiwanica* 26(1). 41-48.
- Sun, F-C., Feng, D.G. (1999) Eye movements in reading Chinese and English text. In Wang, Inhoff and Chen (eds.) (1984). 189-206.
- Takavolian, S. (ed.) (1981) *Language acquisition and linguistic theory*. Cambridge Ma. MIT Press.
- Tan, C-L. (1995) Language teaching and translation. In Chan and Pollard (eds.) (1995) 476-486.
- Tang, S.W. (2000) Characteristics of some verbless sentences in Chinese. Paper given at Linguistic Society of Hong Kong Annual Research Forum 2000.
- Tao H-Y. (1996) *Units in Mandarin conversation: prosody, discourse and grammar*. Amsterdam. John Benjamins Publishing Company.

- Tao, L. (1993) Zero anaphora in Chinese: cognitive strategies in discourse processing. Ph. D. dissertation. Boulder. University of Colorado.
- Tao, L. (1996) Topic discontinuity and zero anaphora in Chinese discourse: cognitive strategies in discourse processing. In Fox (ed.) (1996) 487-514.
- Taylor I. and Olson, D.R. (eds.) (1995) *Scripts and literacy: reading and learning to read alphabets, syllabaries and characters*. Dordrecht. Kluwer Academic Publishers.
- Taylor, I. and Taylor M. (1995) *Writing and literacy in Chinese Korean and Japanese*. Amsterdam. John Benjamins Publishing Company.
- Thorndyke, P.W. (1977) Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology* 9. 135-147.
- Toury, G. (1986) Monitoring discourse transfer: a test case for a developmental model of translation. In House and Blum-Kulka (eds.) (1986).
- Toury, G. (1995) *Descriptive translation studies and beyond*. Amsterdam. John Benjamins Publishing Company.
- Translational English Corpus (2000). UMIST <http://www.umist.ac.uk/ctis/>  
□□<http://www.umist.ac.uk/ctis/>
- Tzeng, O.J.L., Hung, D.L. and Wang, W.S-Y. (1977) Speech recoding in reading Chinese characters. *Journal of experimental psychology: human learning and memory*. 1977. 3. 621-630.
- Tzeng O.J.L. and Hung D.L. (1981) Linguistic determinism: a written language perspective. In Tzeng and Singer (eds.) (1981) 235-255.
- Tzeng, O.J.L. and Singer, H. (eds.) (1981) *Perception of print: reading research in experimental psychology*. Hillsdale, NJ. Lawrence Erlbaum Associates.
- Tzeng, O.J.L., Zhong, H.L., Hung, D. and Lee, W-L. (1995) Learning to be a conspirator: a tale of becoming a good Chinese reader. In de Gelder and Morais (eds.) (1995) 227-246.
- Van Dijk T.A. and Kintsch, W. (1983) *Strategies of discourse comprehension*. New York. Academic Press.
- Vellutino, F. R. And Scanlon, D. M. (1987) Linguistic coding and reading ability. In Rosenberg (ed.) (1987) 1-69.

- Venuti, L. (1995) *The translator's invisibility; a history of translation*. London. Routledge.
- Vinay, J-P and Darbelnet, J. (1958) *Stylistique comparée du Français et de l'Anglais: méthode de traduction*. London. Harrap.
- Waggot V. (1997) Magical re-arrangements: how one hundred bilingual teenagers read three Chinese texts. M. Sc. Dissertation. Department of Applied Linguistics. University of Edinburgh.
- Waggot. V. (2000) Features, cobwebs or clines: towards a possible model of lexical retrieval in bilingual readers. *Edinburgh working papers in applied linguistics*. 10. 72-79.
- Wang, J., Inhoff, A.W. and Chen, H.C. (1999) *Reading Chinese script- a cognitive analysis*. Mahwah, NJ. Lawrence Erlbaum Associates.
- Widdowson, H.G. (1978) *Teaching language as communication*. Oxford. OUP.
- Widdowson, H.G. (1984) *Explorations in Applied Linguistics 2*. Oxford. OUP.
- Winston P. H. (ed.) (1975) *The psychology of computer vision*. New York. McGraw-Hill.
- Wong, C.H. (1992) Cantonese-speaking children's understanding of anaphora. In Lee (1992) 100-159.
- Wong, I. P-K. (1995) Translating empty words. In Chan and Pollard (1995) 209-215.
- Woodworth, R.S. (1938) *Experimental Psychology*. New York. Holt.
- Wu, J.T. and Liu, I.M. (1988) A data base system about the psychological features of Chinese characters and words. In Liu, Chen and Chen (eds.) (1998) 171-186.
- Xu Suchong (1984) *Fanyi de yishu*. Beijing. Zongguo Duiwai Fanyi Chubanshe.
- Xu J-J. (1990) Shushu wenzhong 'ta' de huayu fenxi. *Zhongguo Yuwen* 1990. No 5. 325-337.
- Xu S-Y. (1991) Linghuoxing he yuanzexing - jiantan yiwei he shenglue. *Zhongguo Yuyanxuebao*. Vol. 6. 62-70.

Yan, J.X. (2000) Language teaching in an English-Putonghua interpretation course. Paper given at Linguistic Society of Hong Kong Annual Research Forum 2000.

Yang, C.L. Gordon, P.C., Hendrick, R. and Wu, J.T. (1999) Comprehension of referring expressions in Chinese. In Chen and Zhou (eds.) (1999) 715-743.

Yang, H-M. and McConkie, G.W. (1999) Reading Chinese: some basic eye-movement characteristics. In Wang, Inhoff and Chen (eds.) (1999) 207-222.

Young, L.W-L. (1994) *Cross-talk and culture in Sino-American communication*. Cambridge. CUP.

Yu, D. (2000) A proposal to set up Chinese-English translation skills courses in secondary schools. (Jianyi zai zhongxueli sheli zhongyiyiing jiqiao yike.) Paper given at International Language in Education Conference 2000. Hong Kong.

Zhang G-J. and Simon, H.A. (1985) STM capacity for Chinese words and idioms: chunking and acoustical loop hypothesis. *Memory and Cognition* 1985 13(3). 193-201.

Zhang G-X. (1993) Tan yinhan. *Zhongguo Yuwen* 1993 No. 2. 126-133.

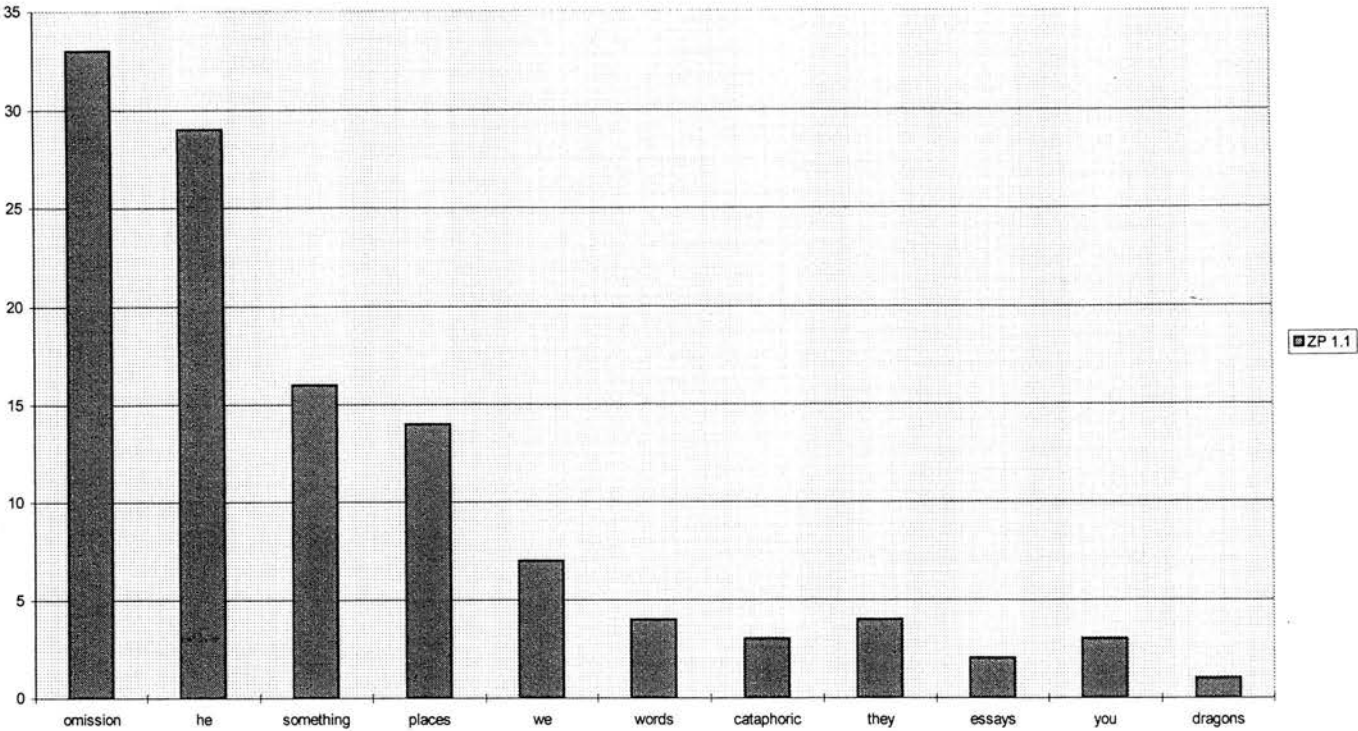
Zhao X-H. (2000) On textual expansion rules; with special reference to ads in Chinese. Paper given at XIII European Association of Chinese studies Conference. Turin 2000.

## Appendices

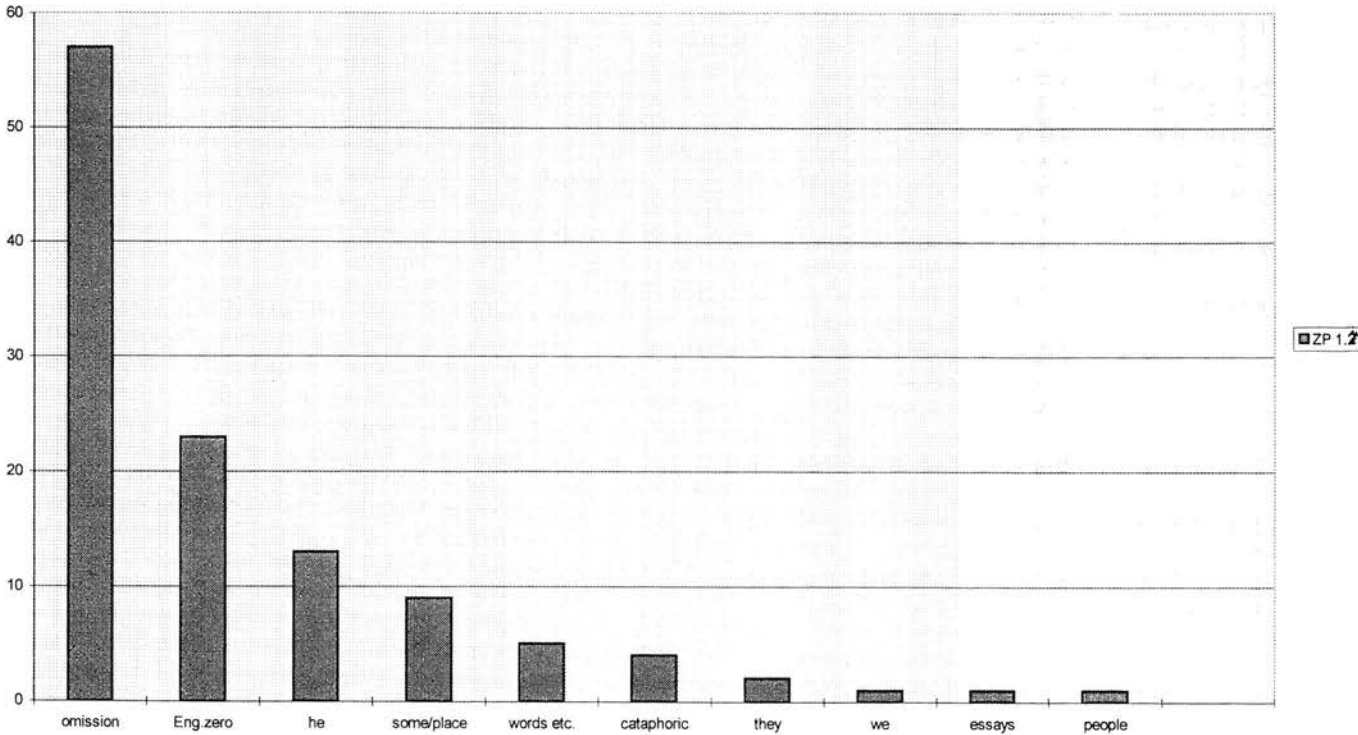


**Appendix I: tables of subjects' actual anaphoric resolution**

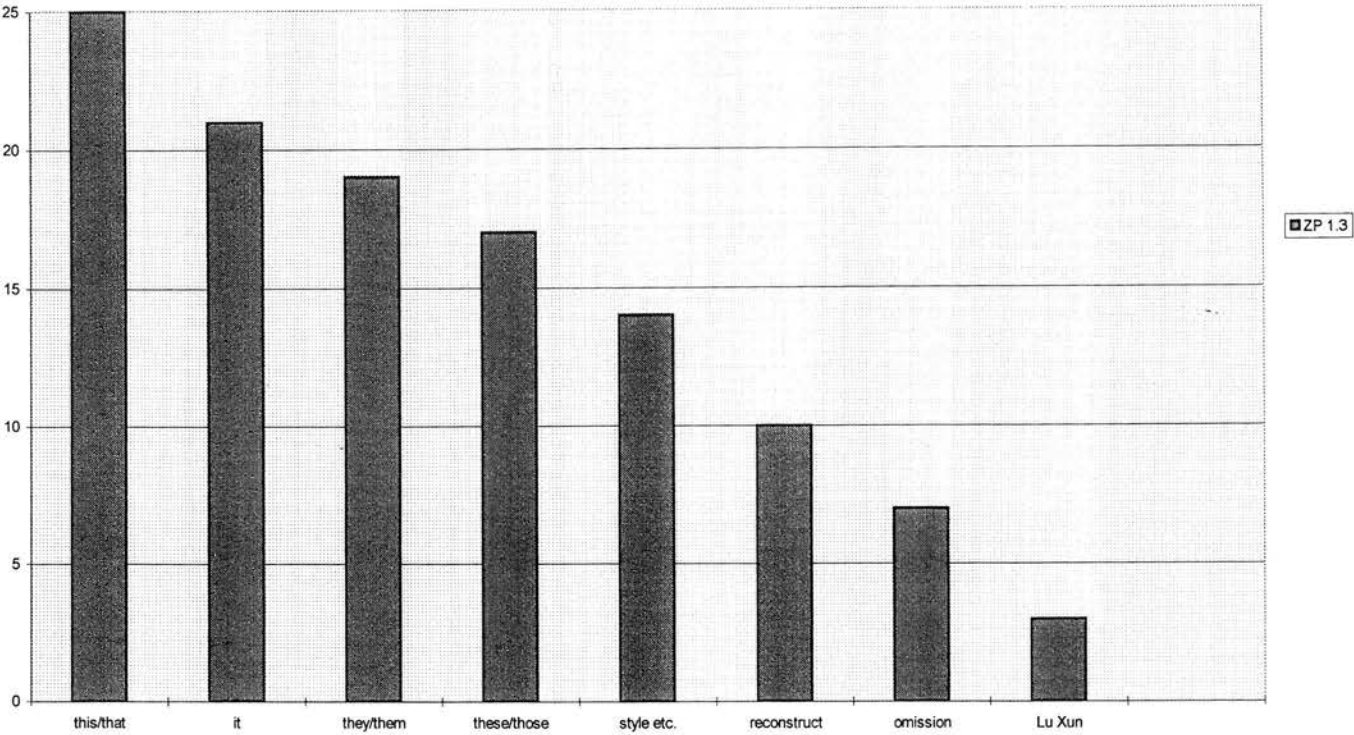
Interpretation of ZP 1.1



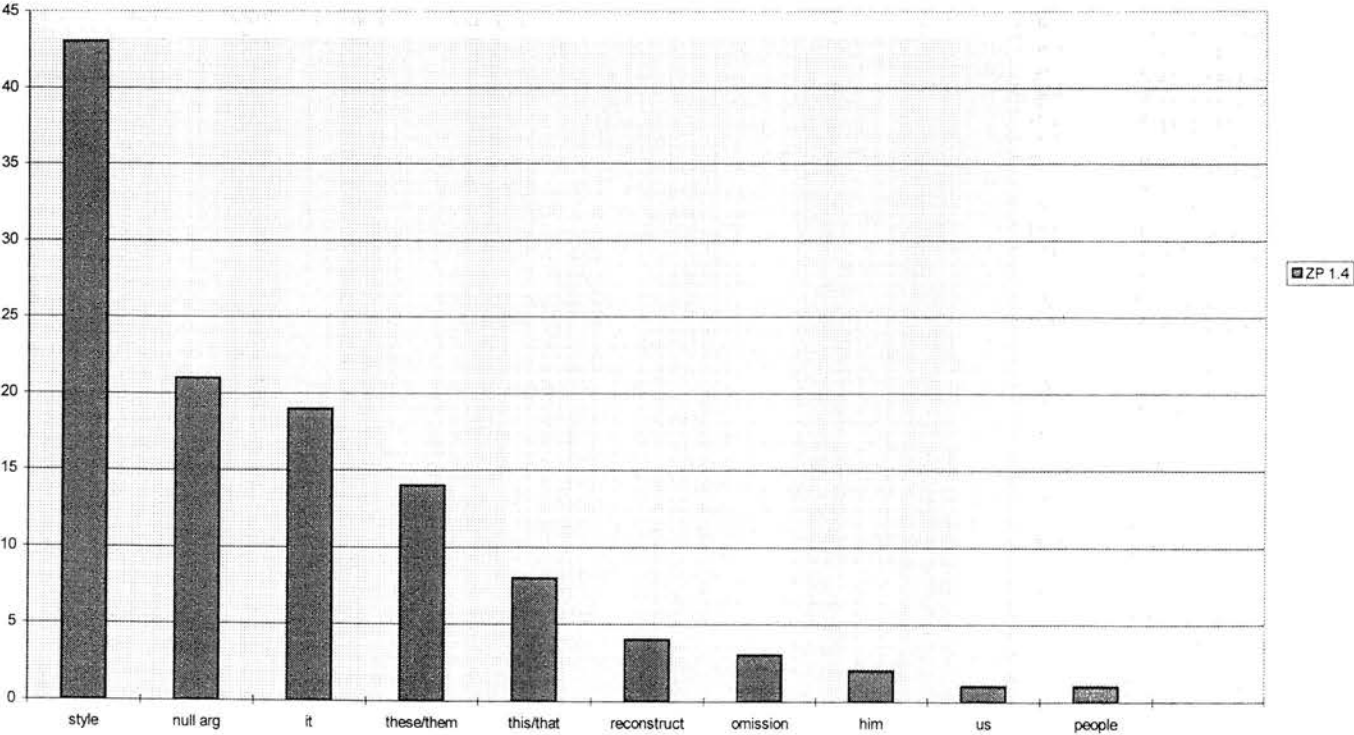
Interpretation of ZP 1.2



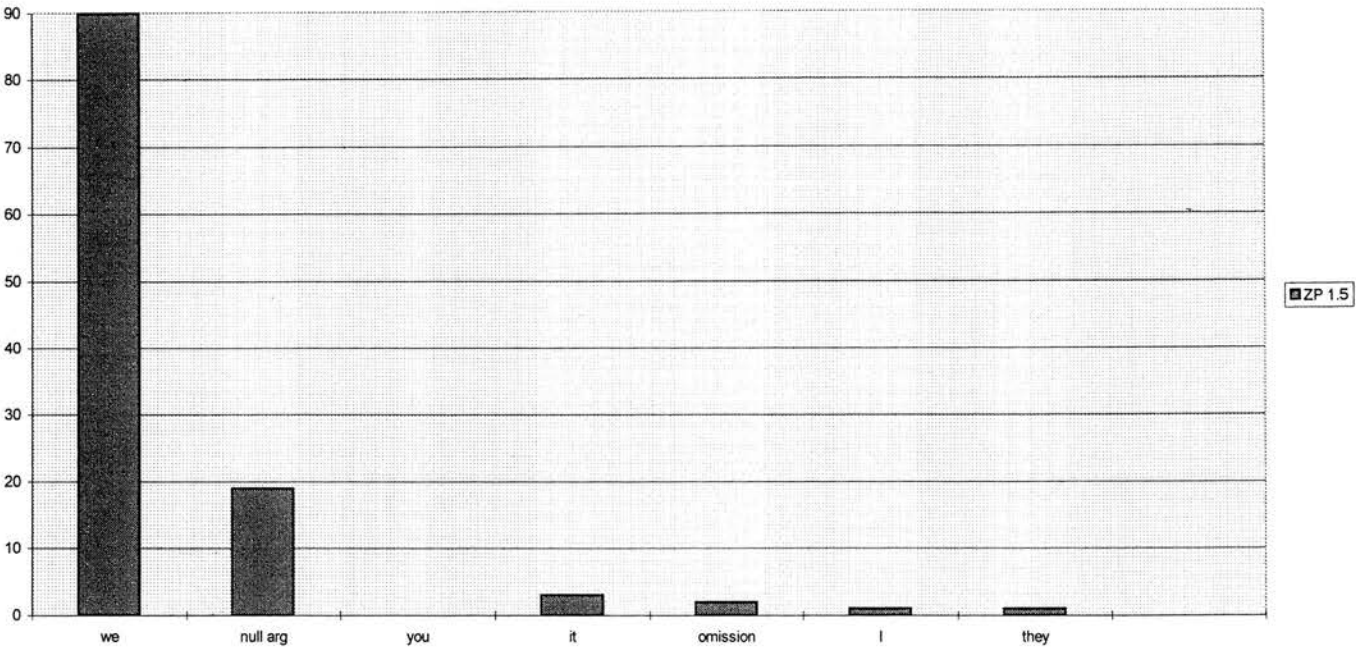
Interpretation of ZP 1.3



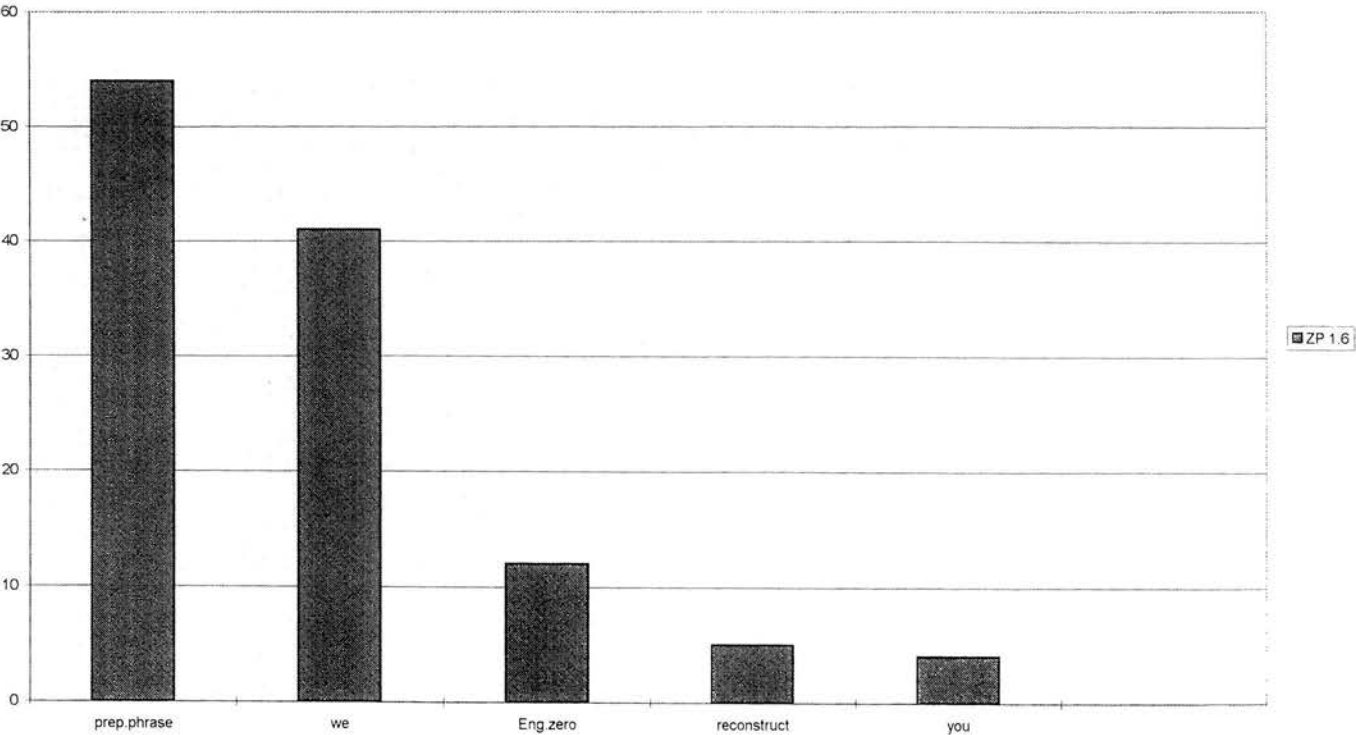
Interpretation of ZP 1.4



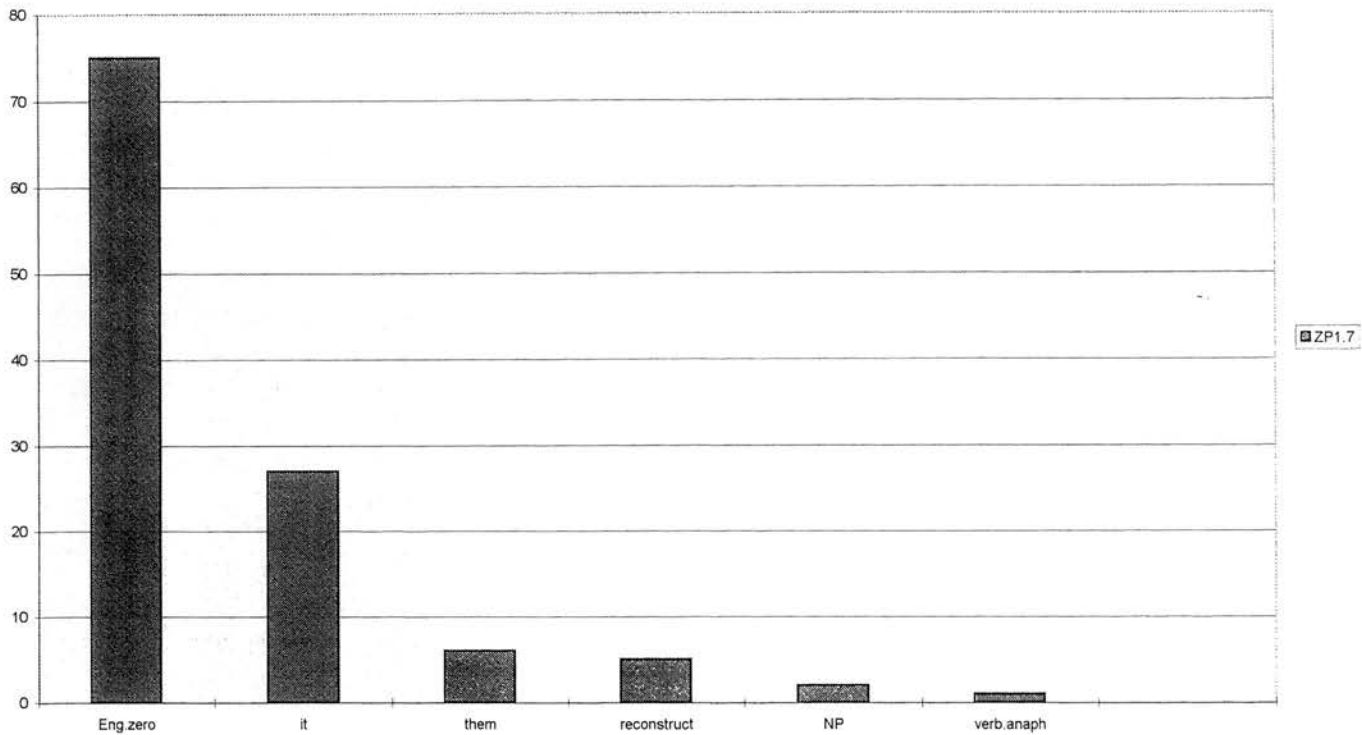
Interpretation of ZP 1.5



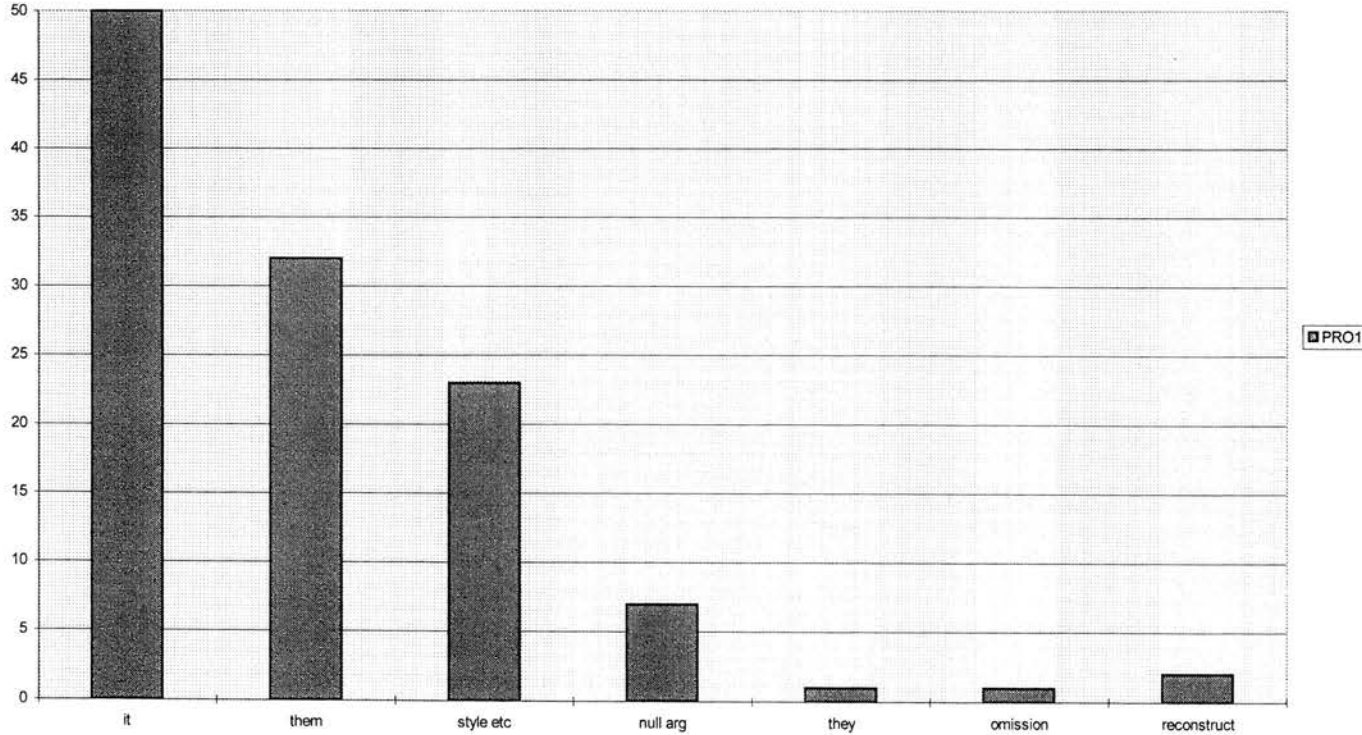
Interpretation of ZP 1.6



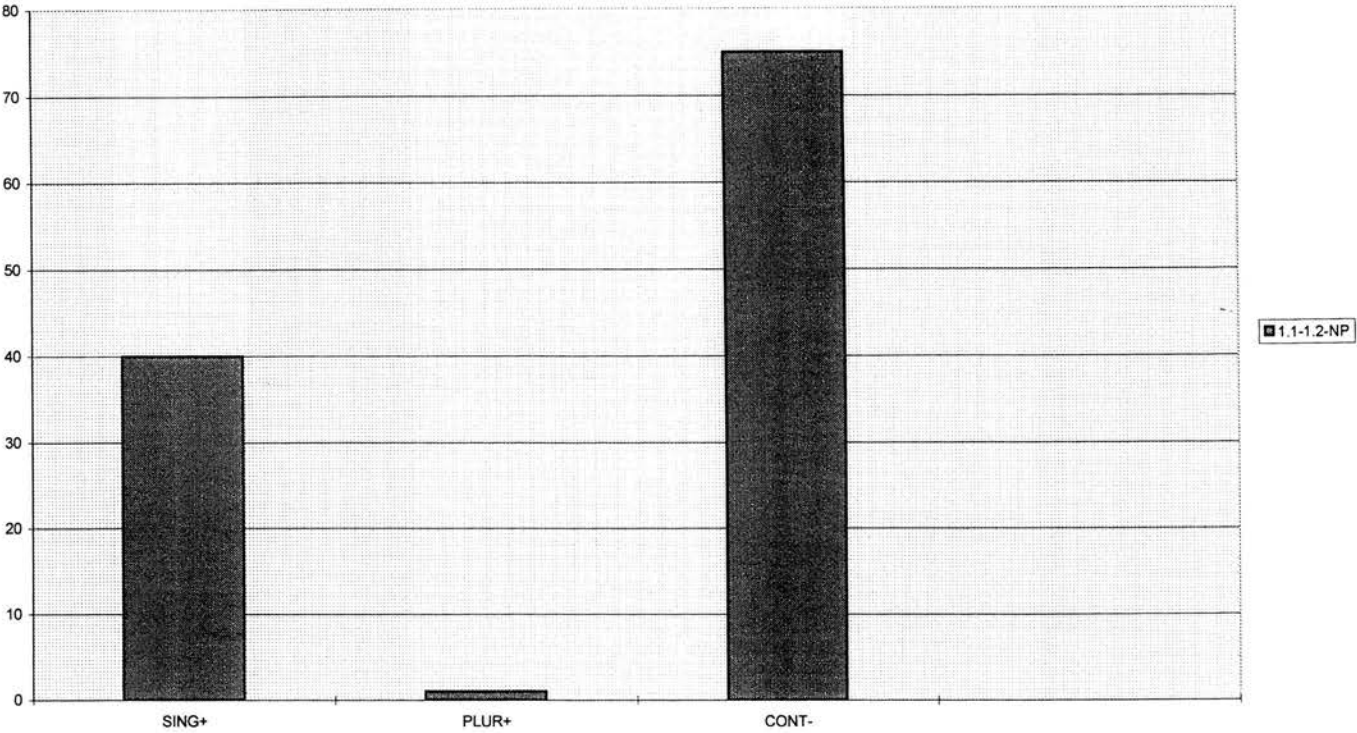
Interpretation of ZP1.7



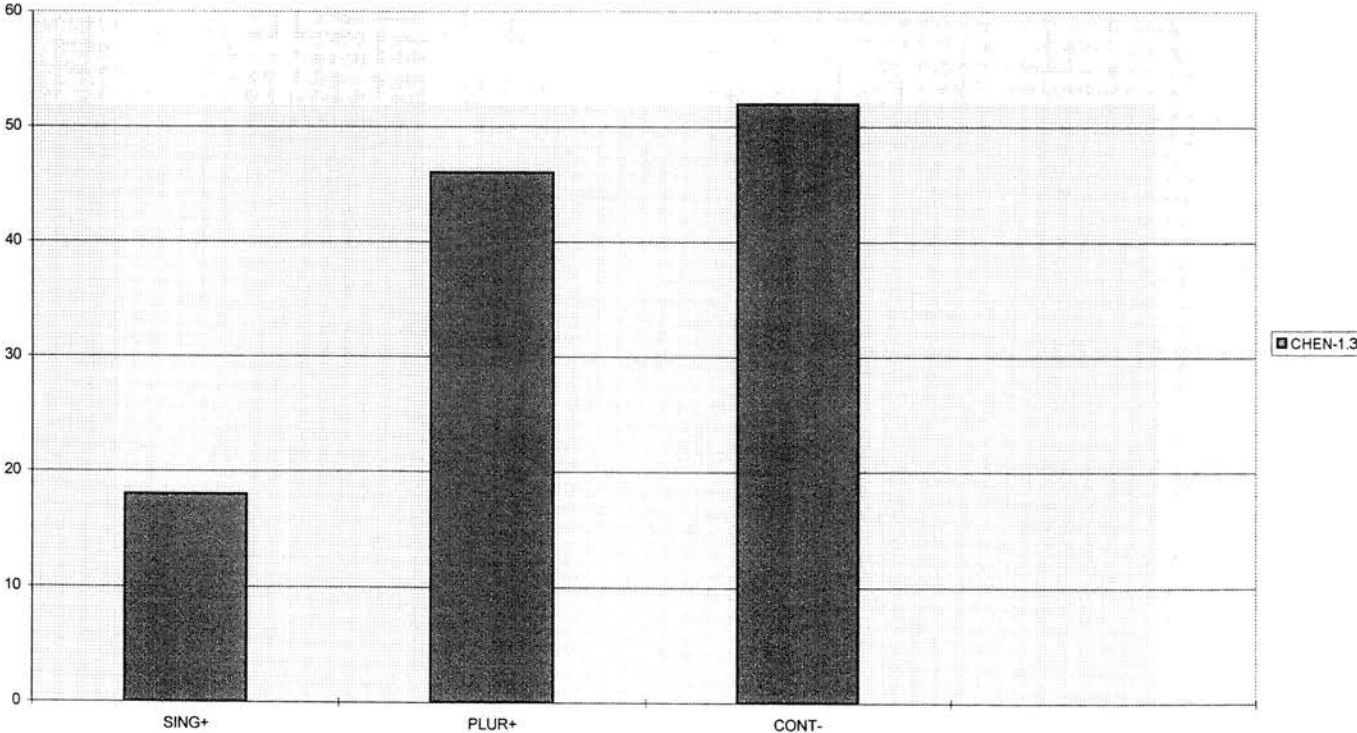
Interpretation of PRO 1



Continuity of Reference ZP 1.1-ZP1.2-NP1

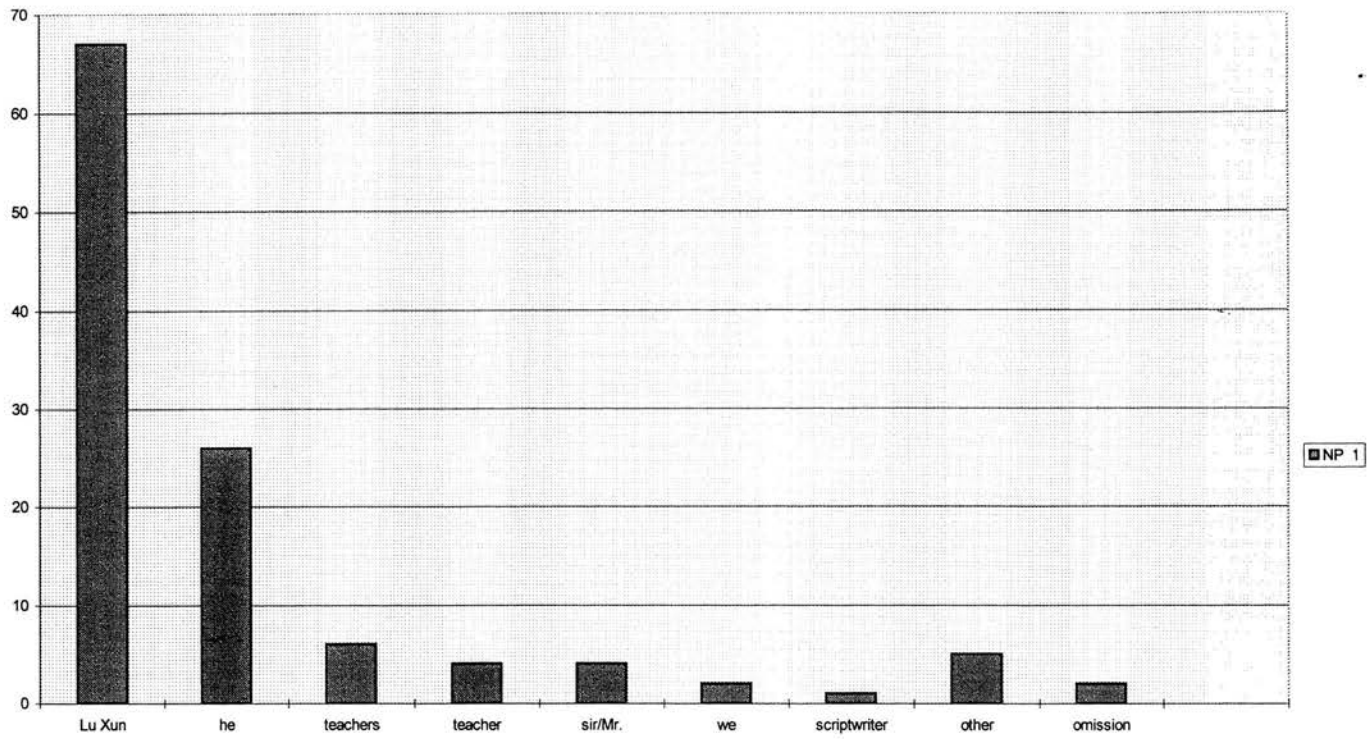


Continuity of reference Chengyu - ZP1.3

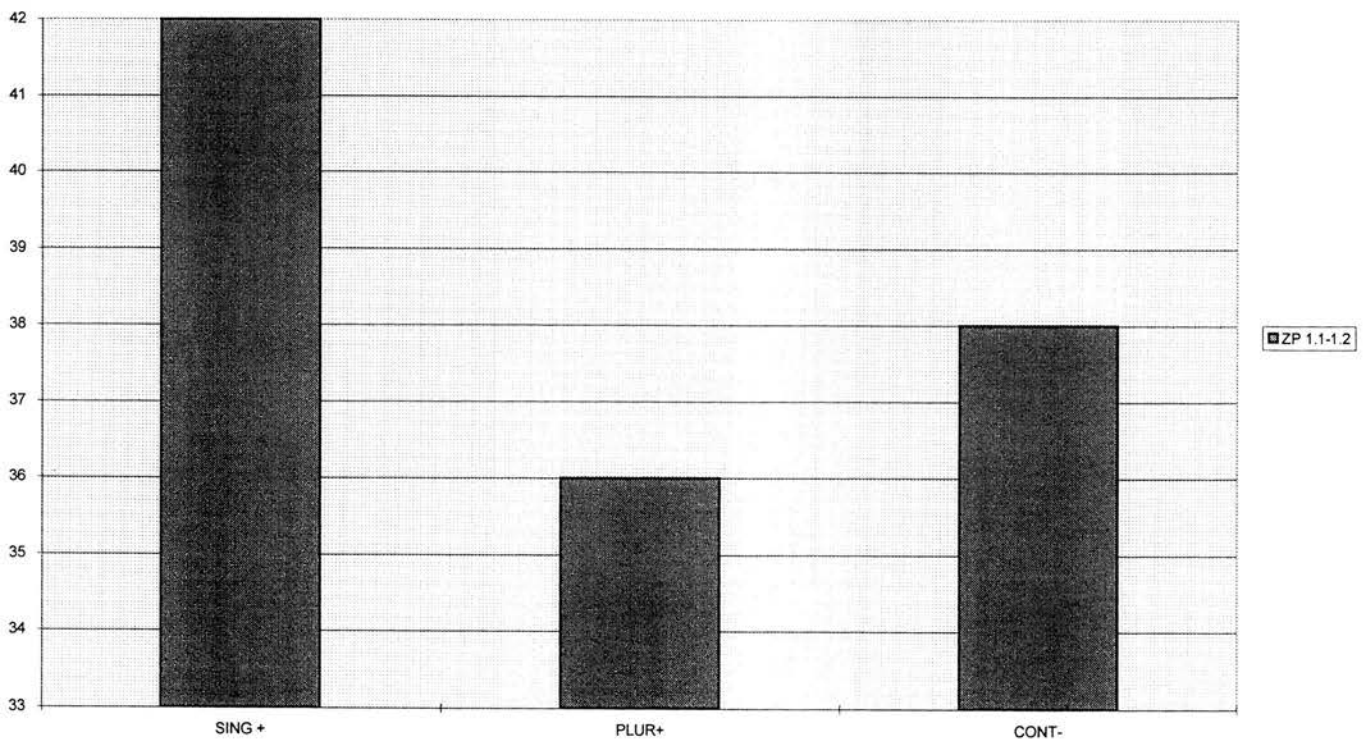




Interpretation of NP 1

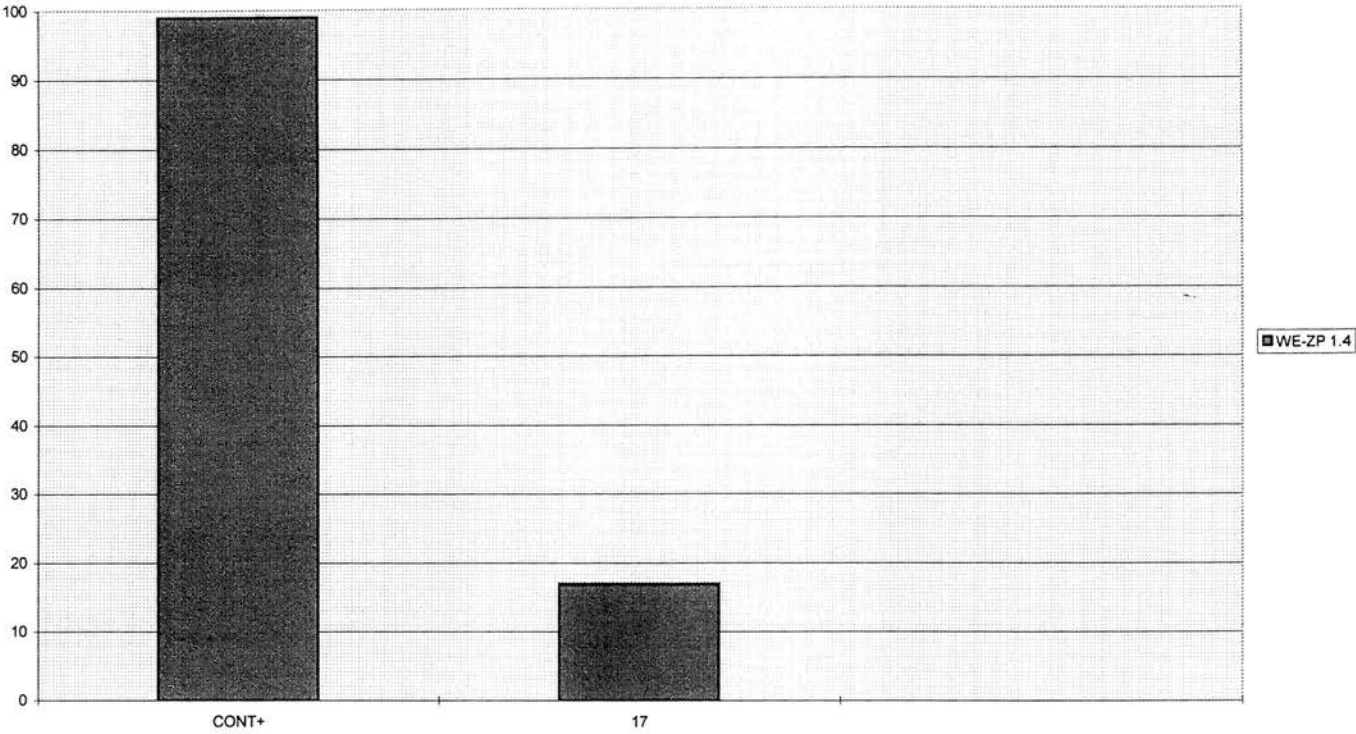


Continuity of reference ZP 1.1 - ZP 1.2

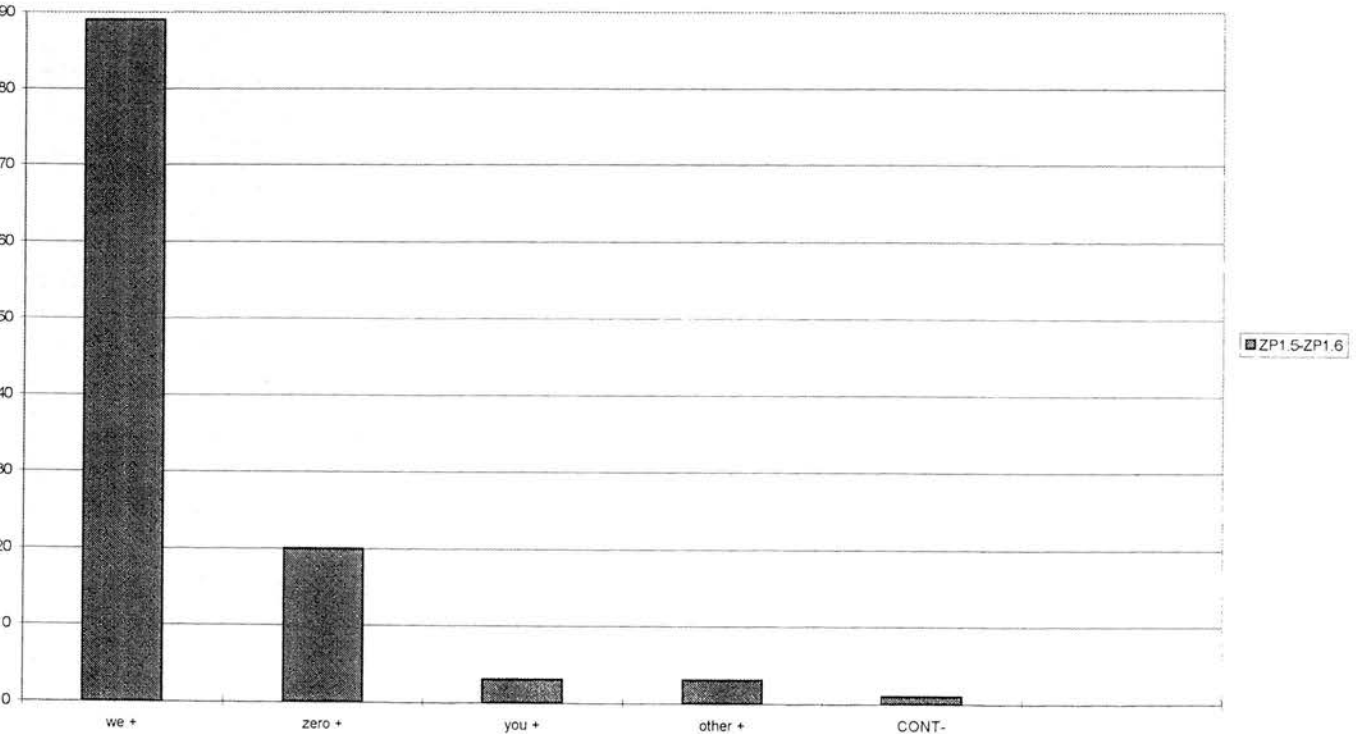




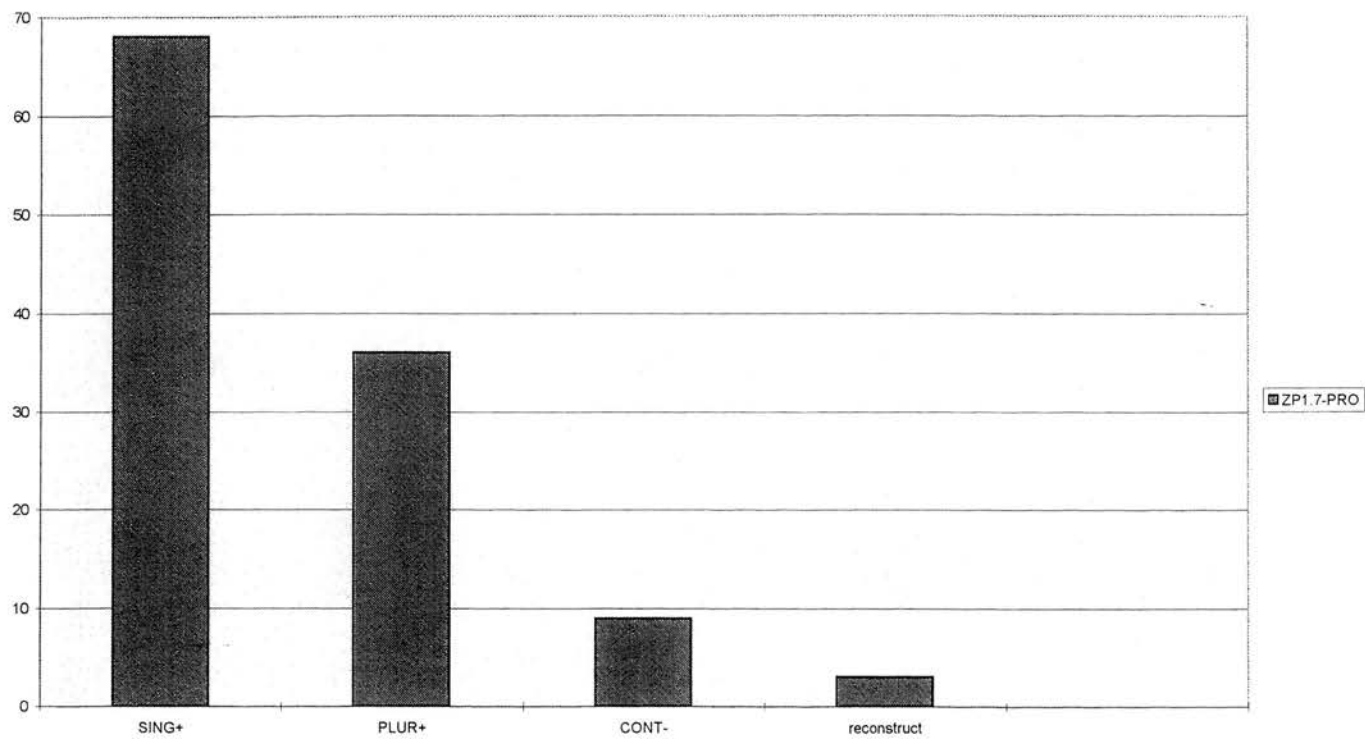
Continuity of reference 'WE' - ZP 1.4 - ZP 1.5



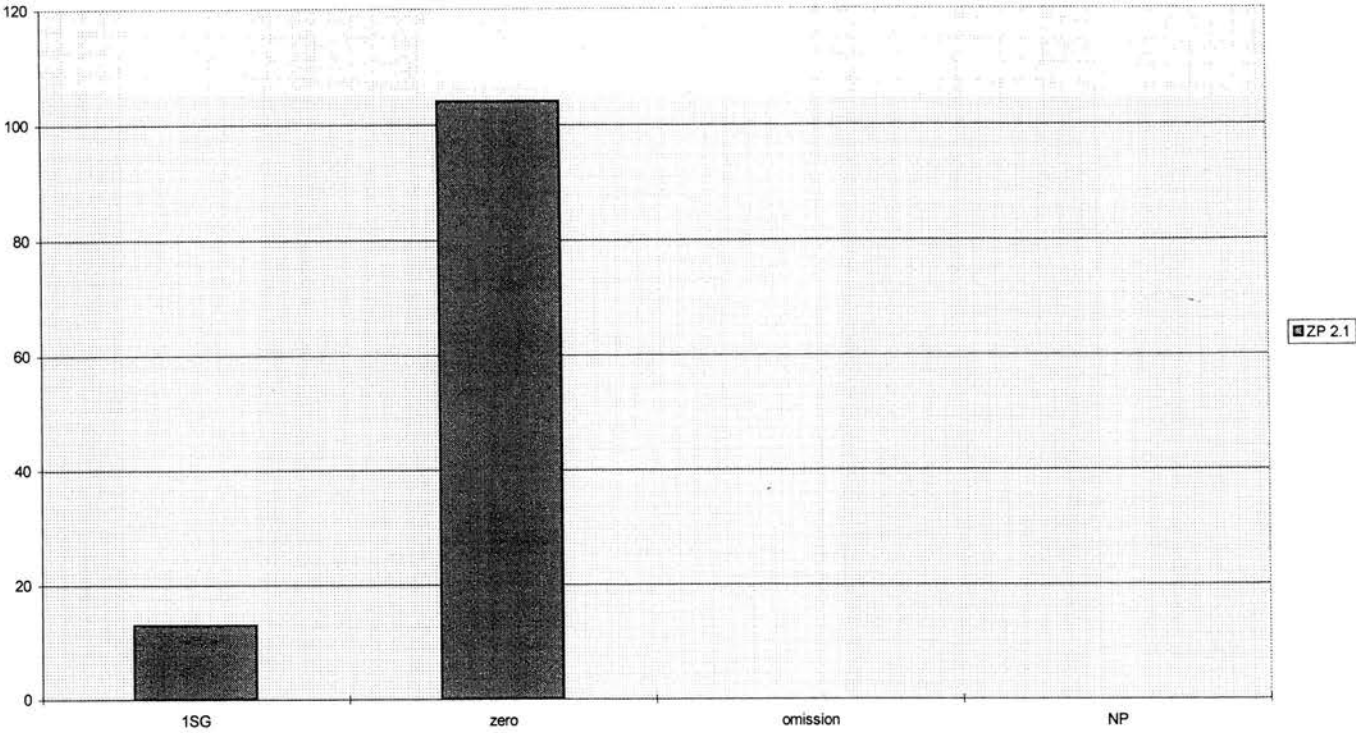
Continuity of reference ZP 1.5 - ZP 1.6



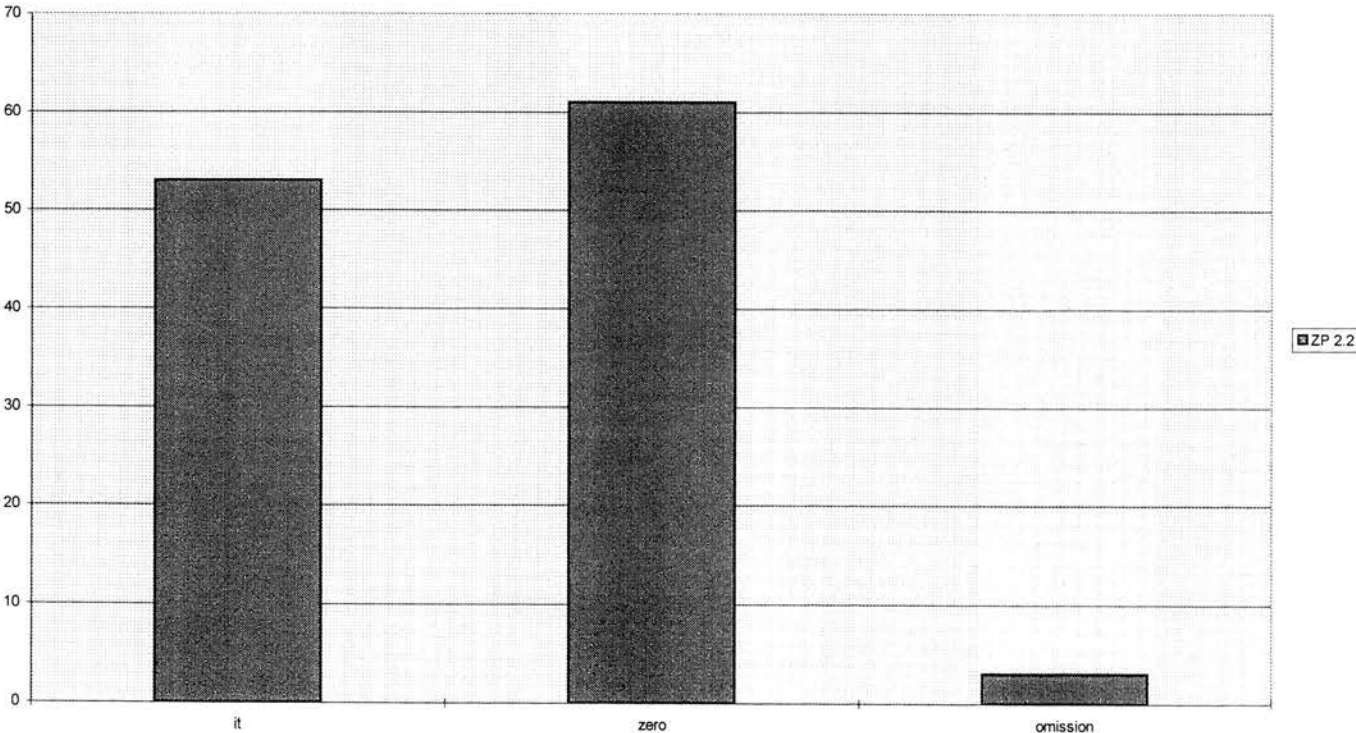
Continuity of reference ZP 1.7 - PRO1



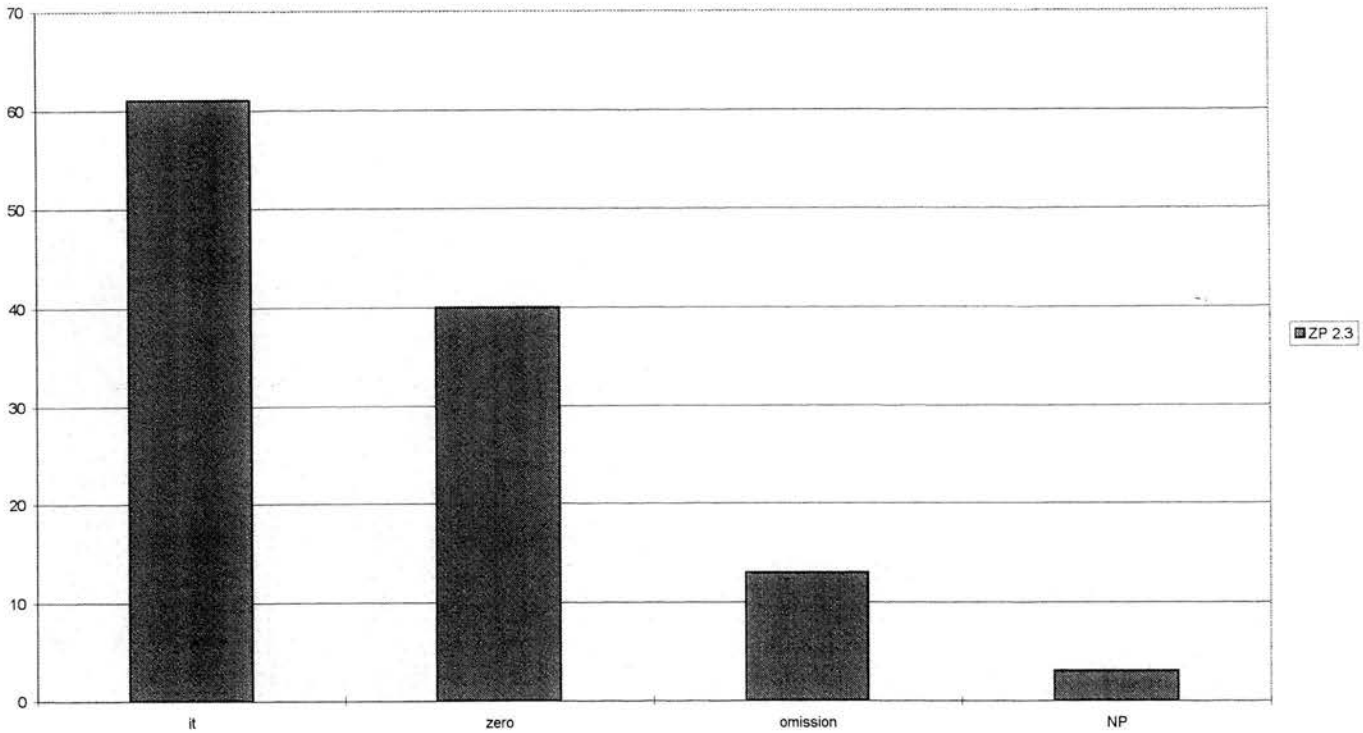
Interpretation of ZP 2.1



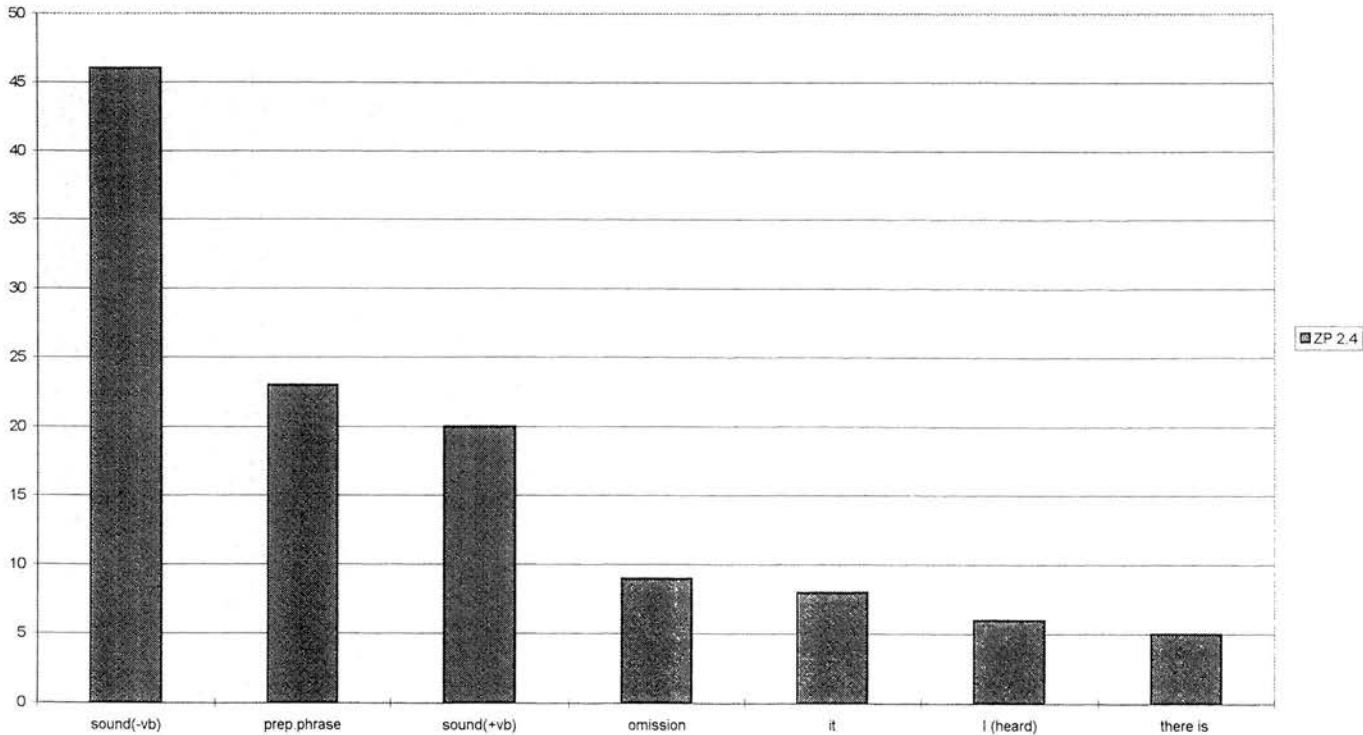
Interpretation of ZP 2.2



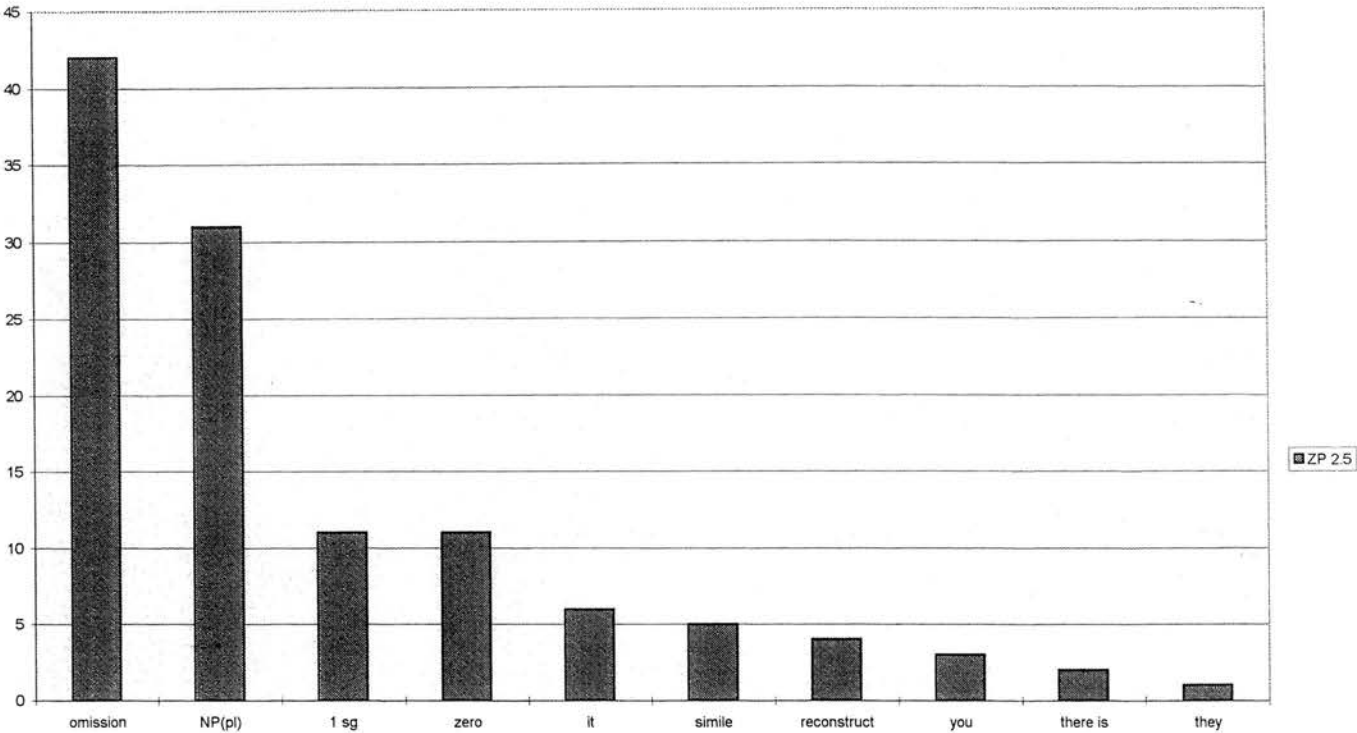
Interpretation of ZP 2.3



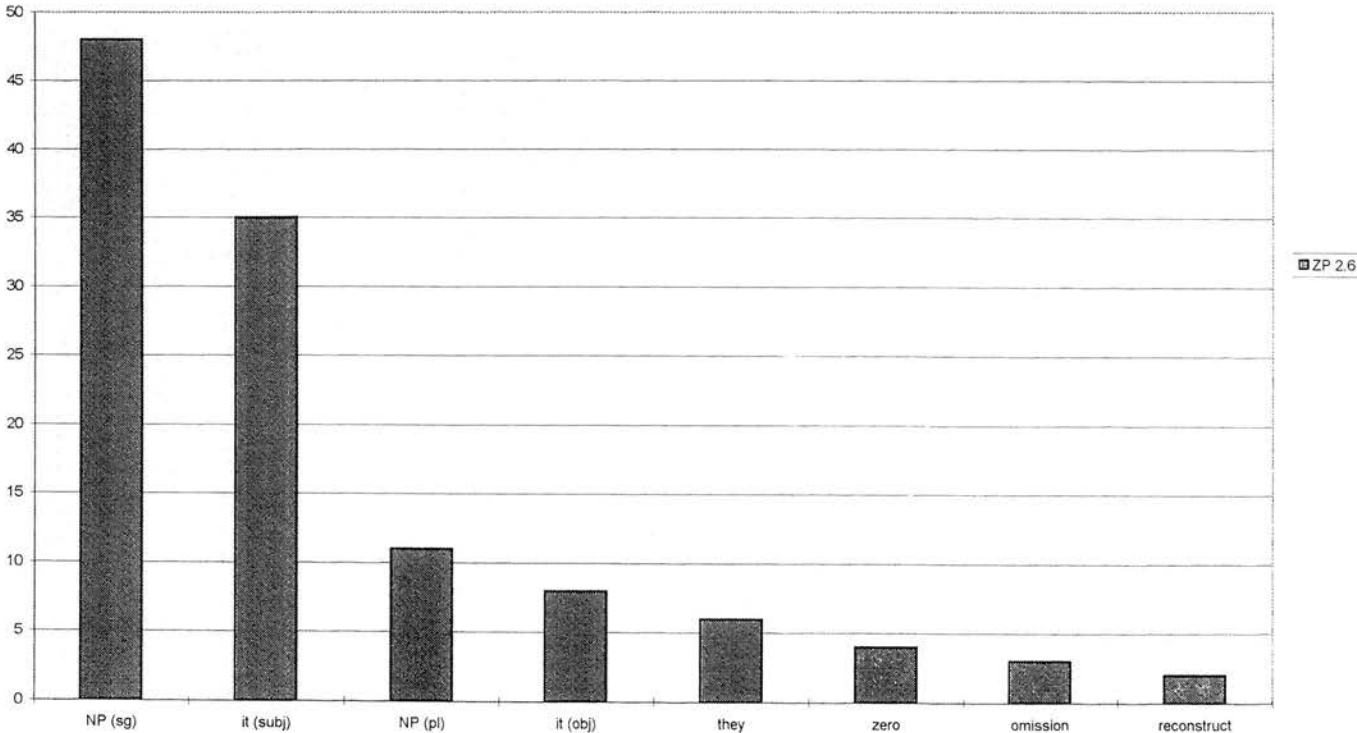
Interpretation of ZP 2.4



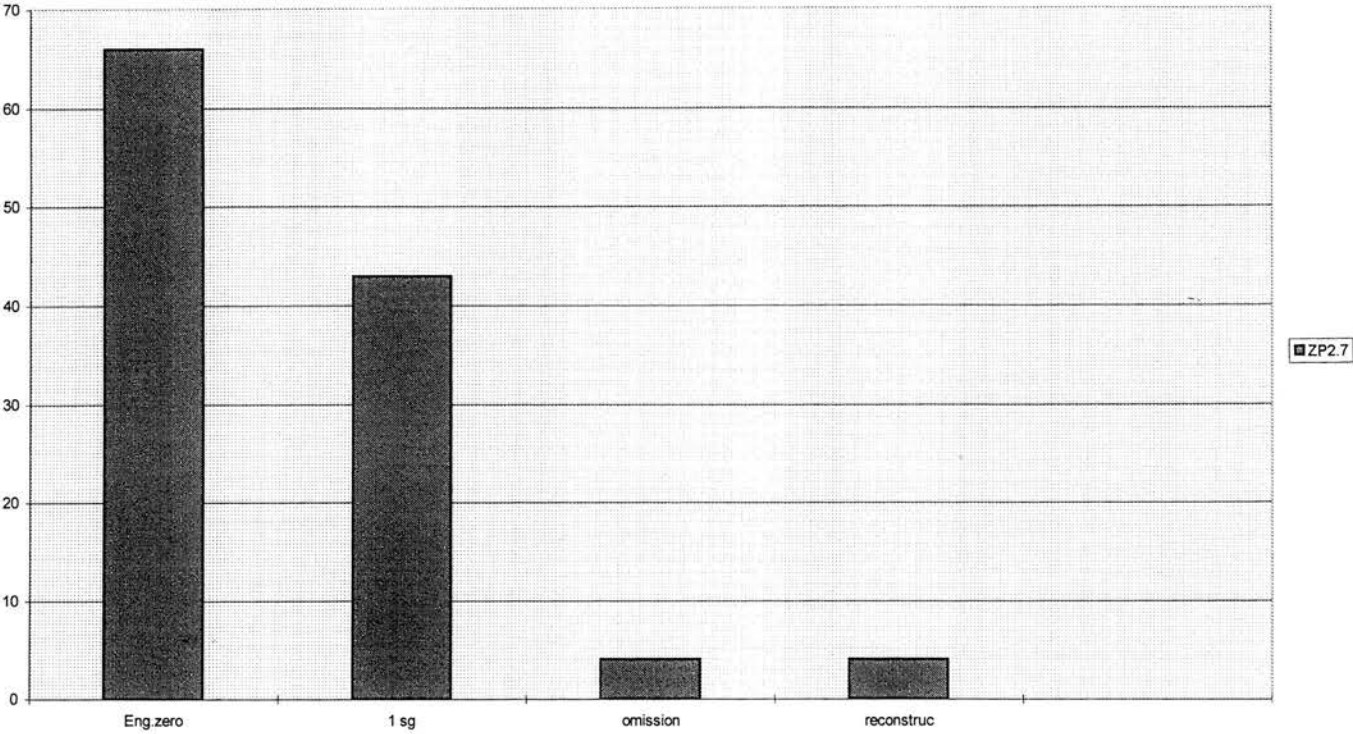
Interpretation of ZP 2.5



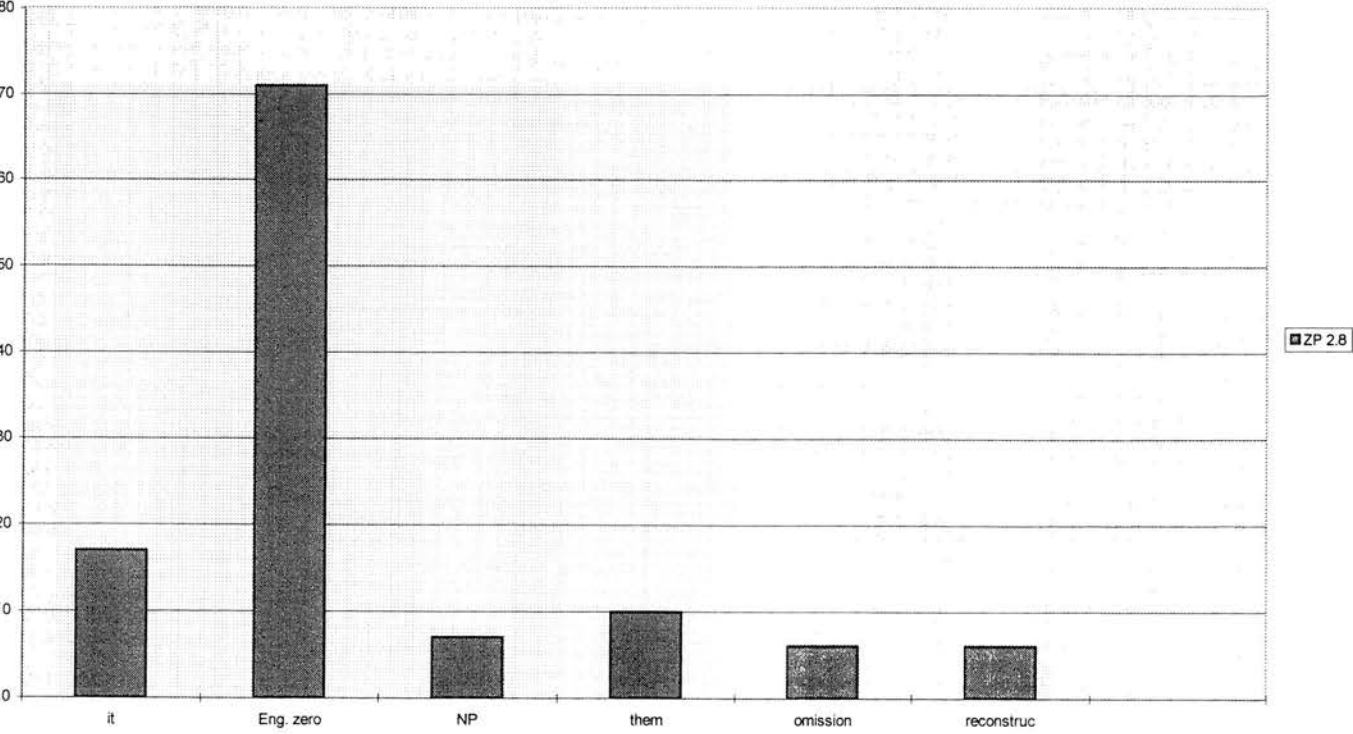
Interpretation of ZP 2.6



Interpretation of ZP 2.7

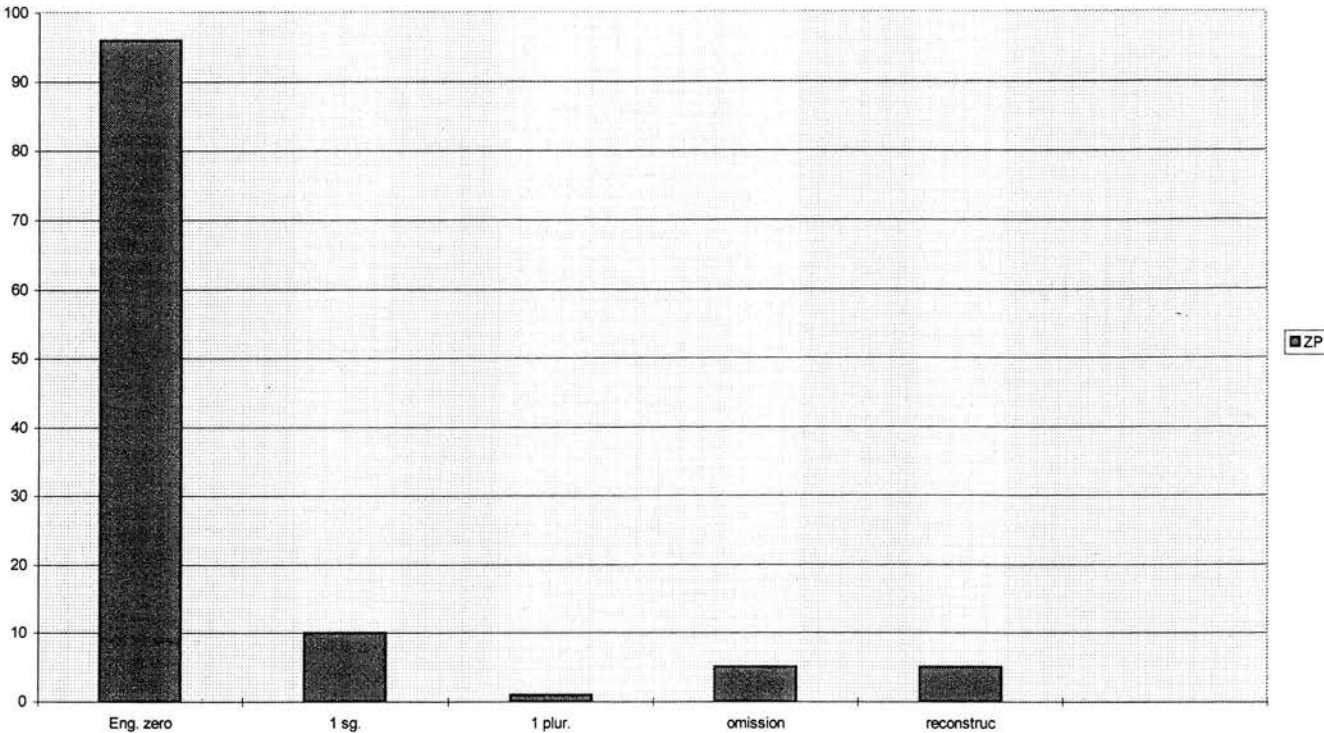


Interpretation of ZP 2.8

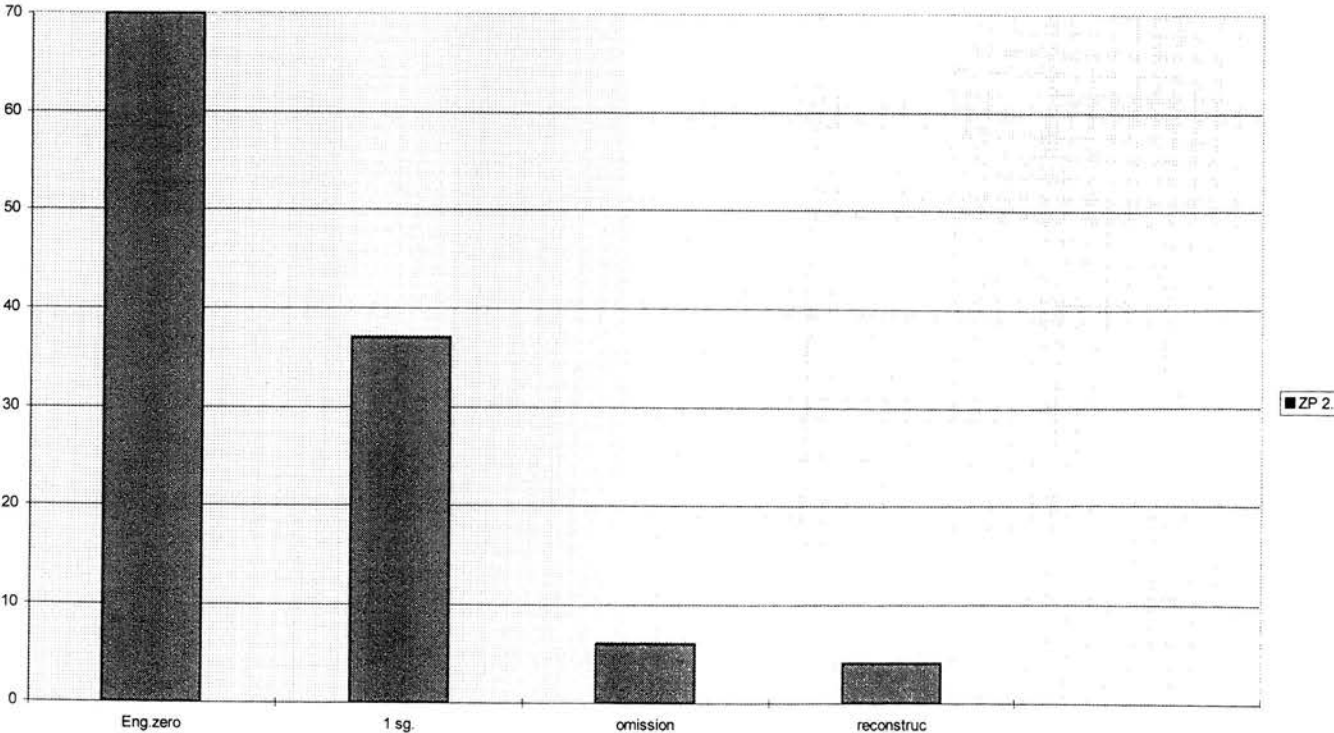




Interpretation of ZP 2.9

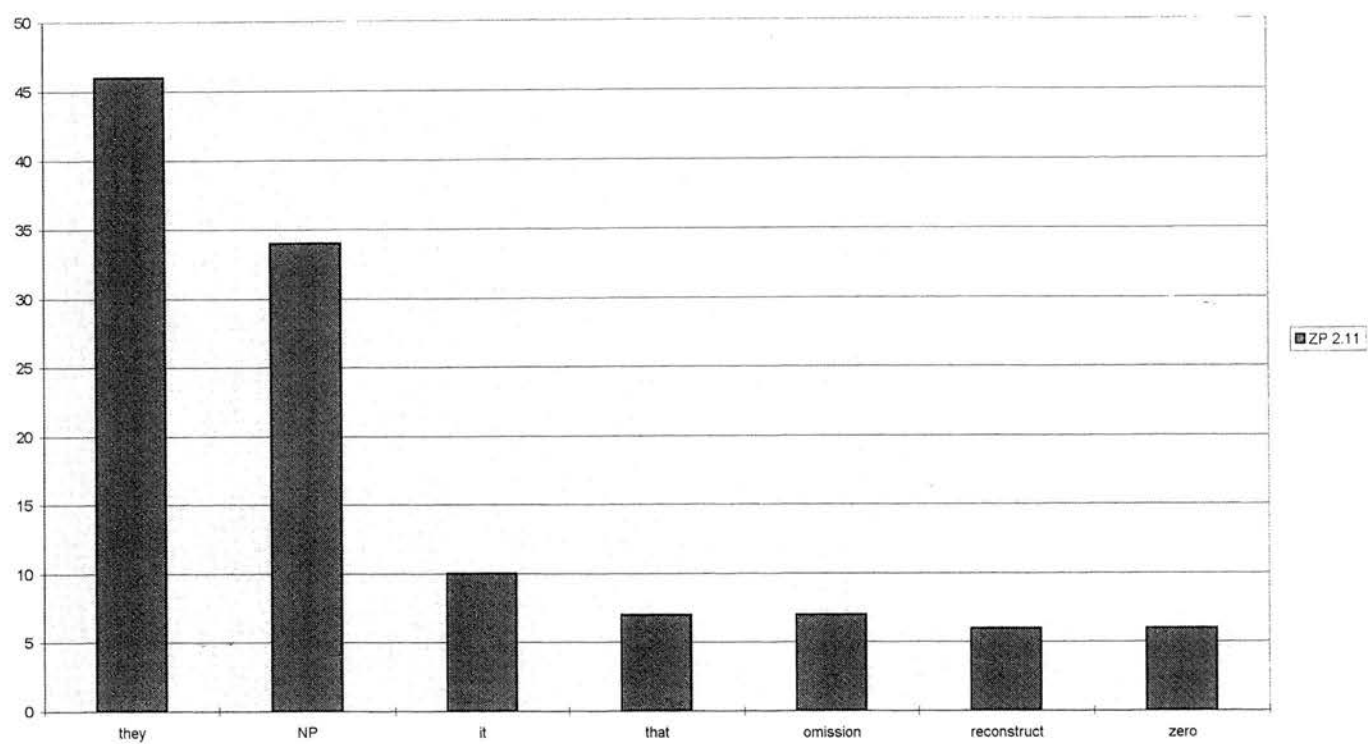


Interpretation of ZP 2.10

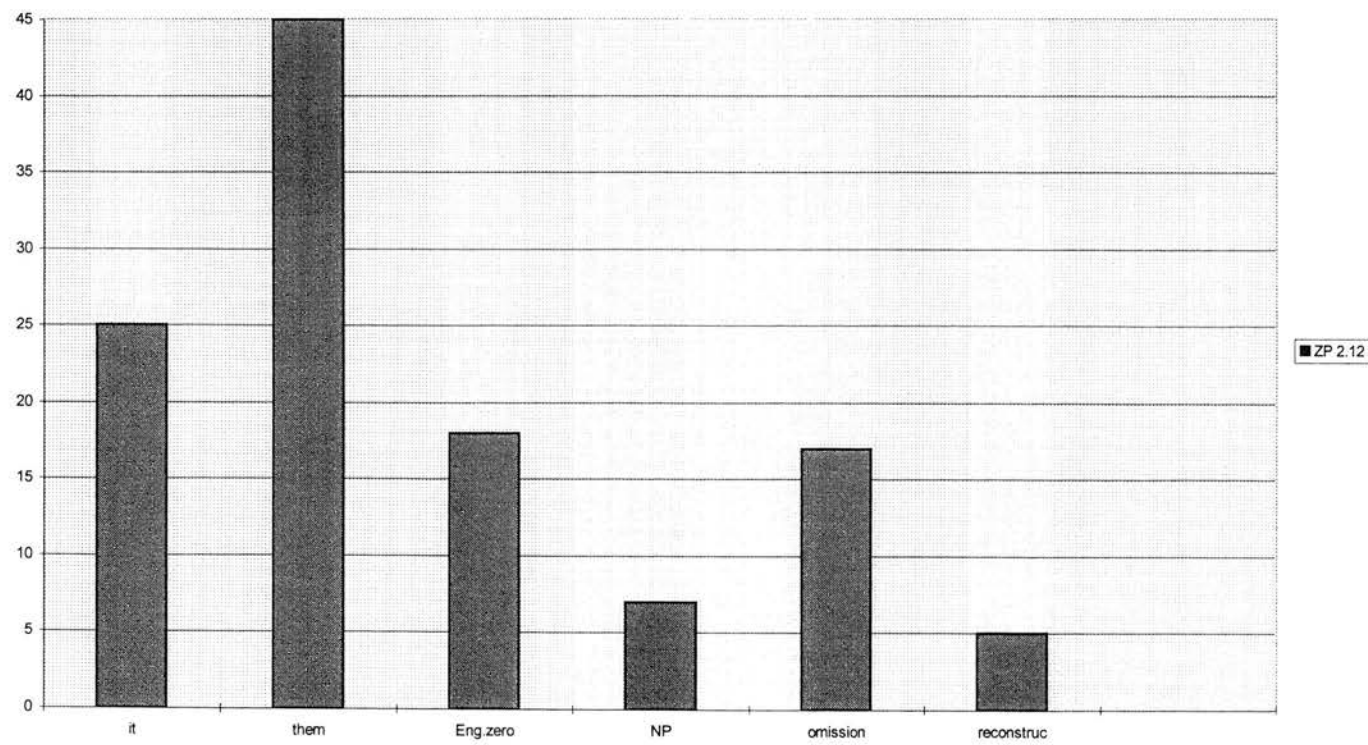




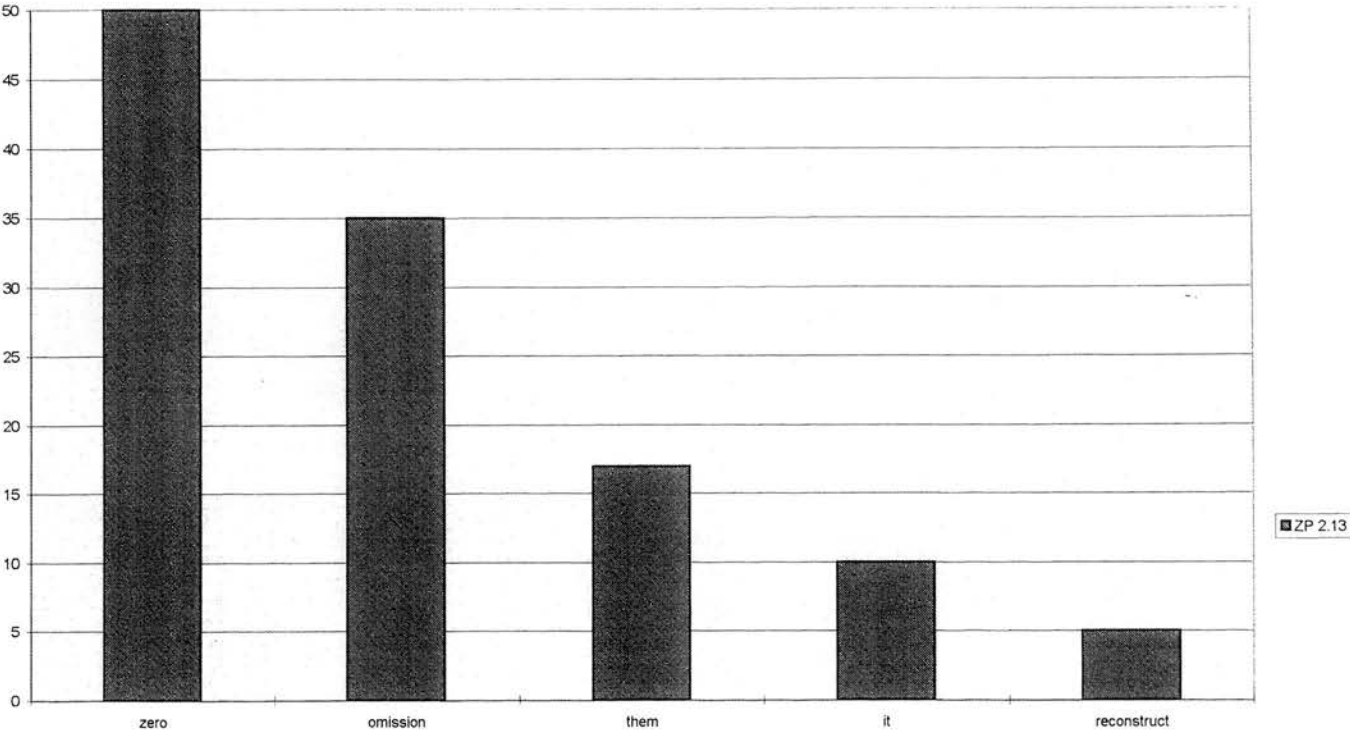
Interpretation of ZP 2.11



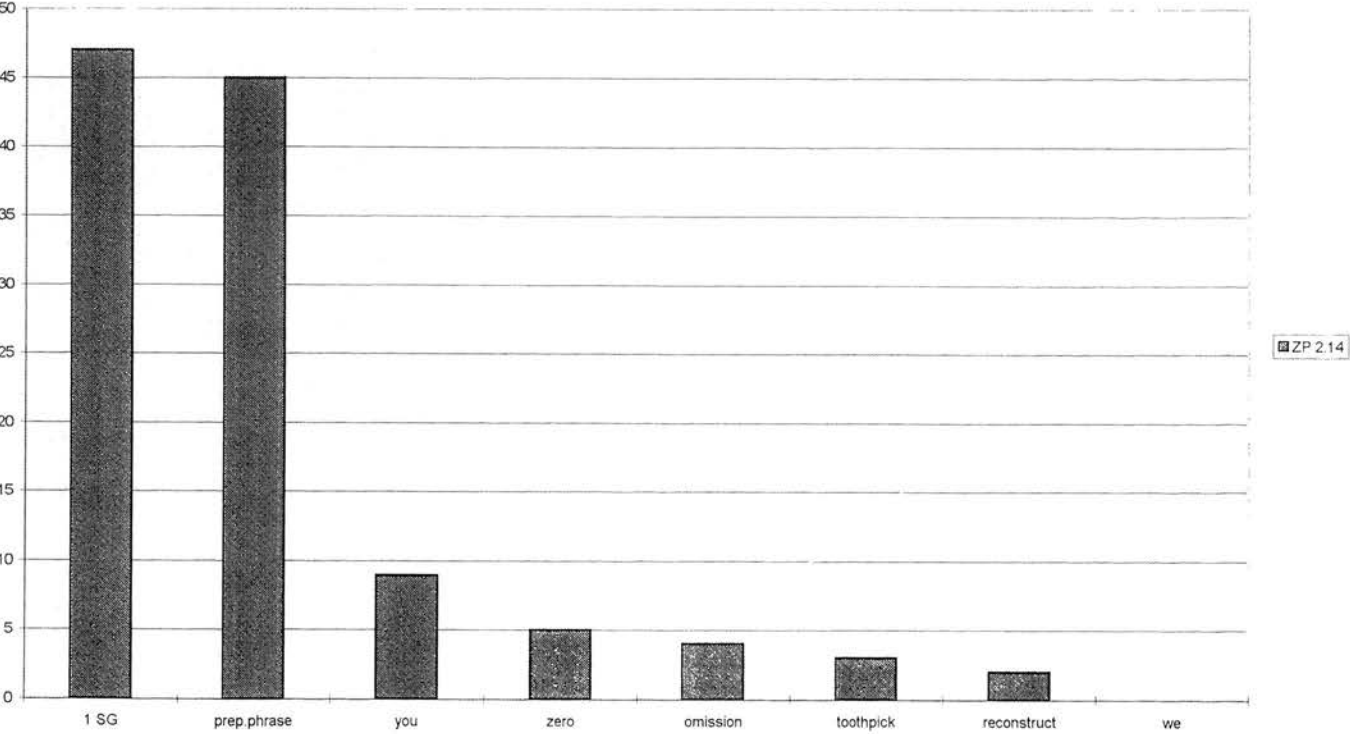
Interpretation of ZP 2.12



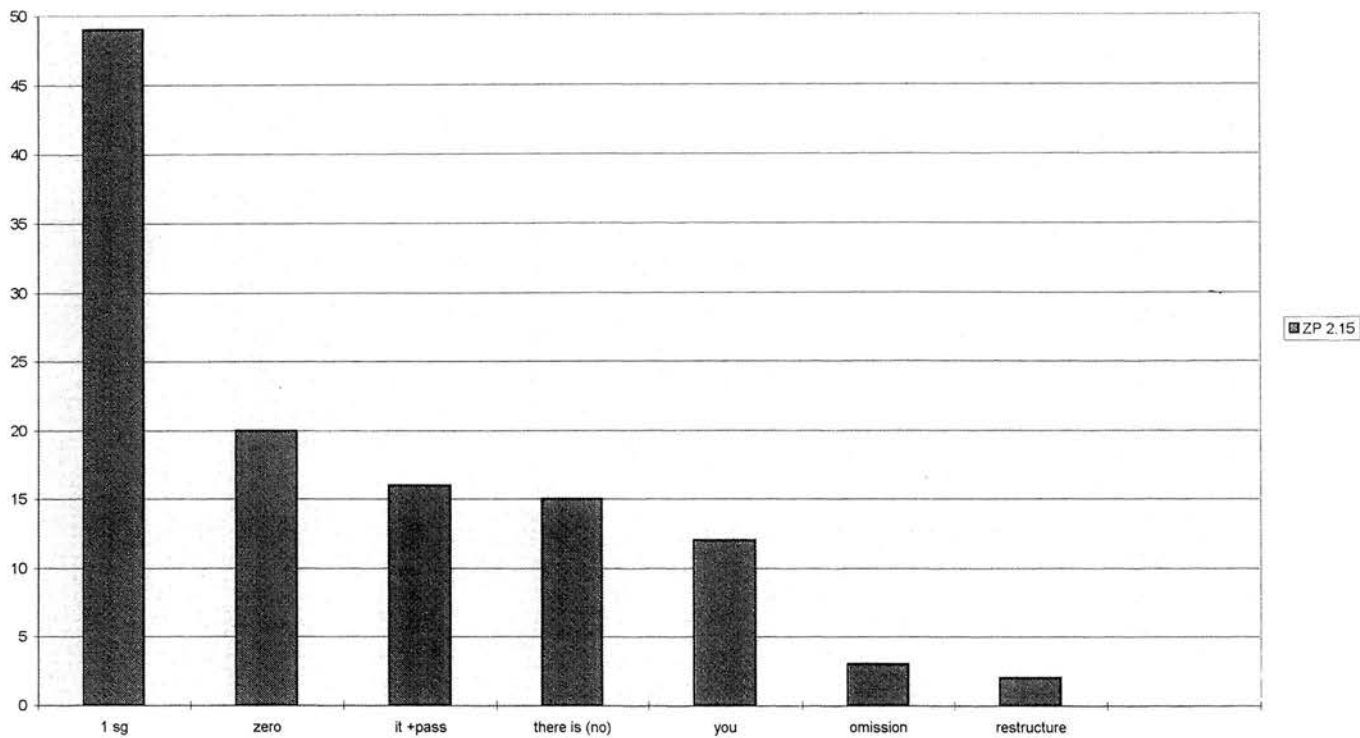
Interpretation of ZP 2.13



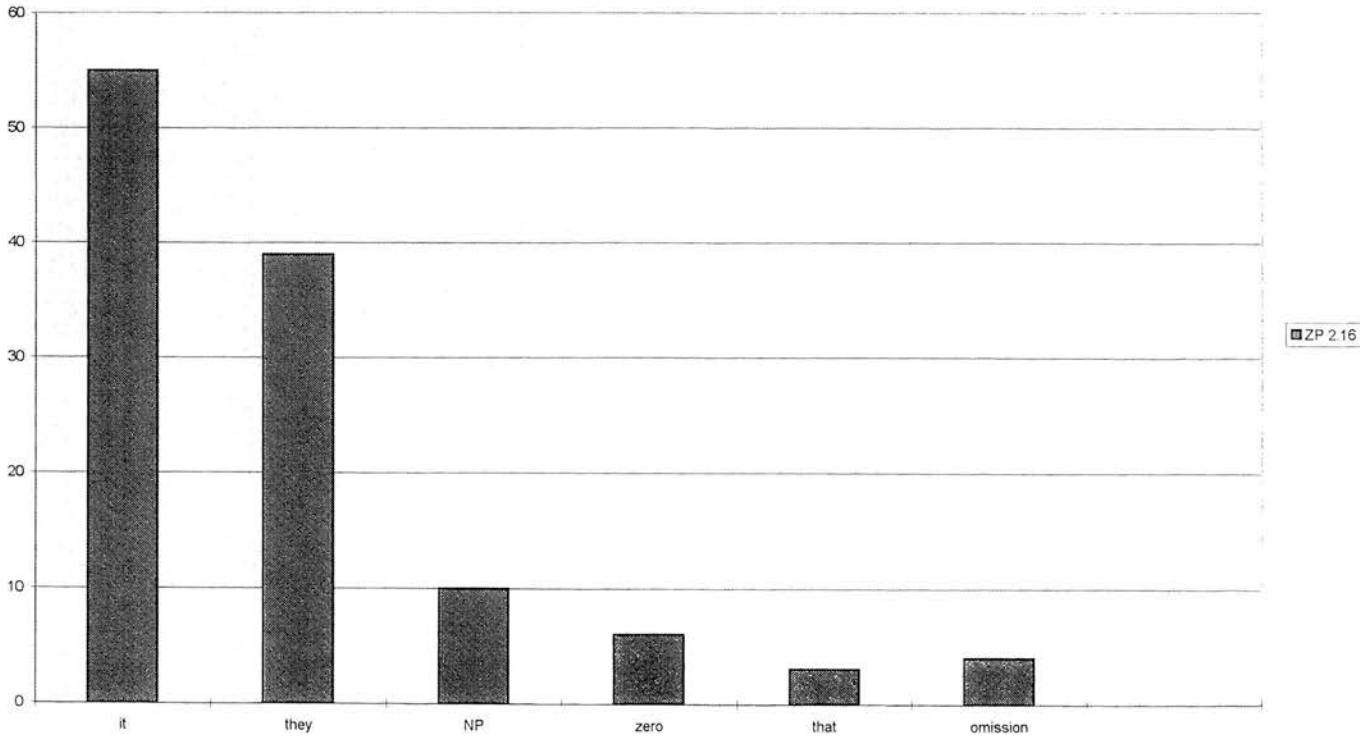
Interpretation of ZP 2.14



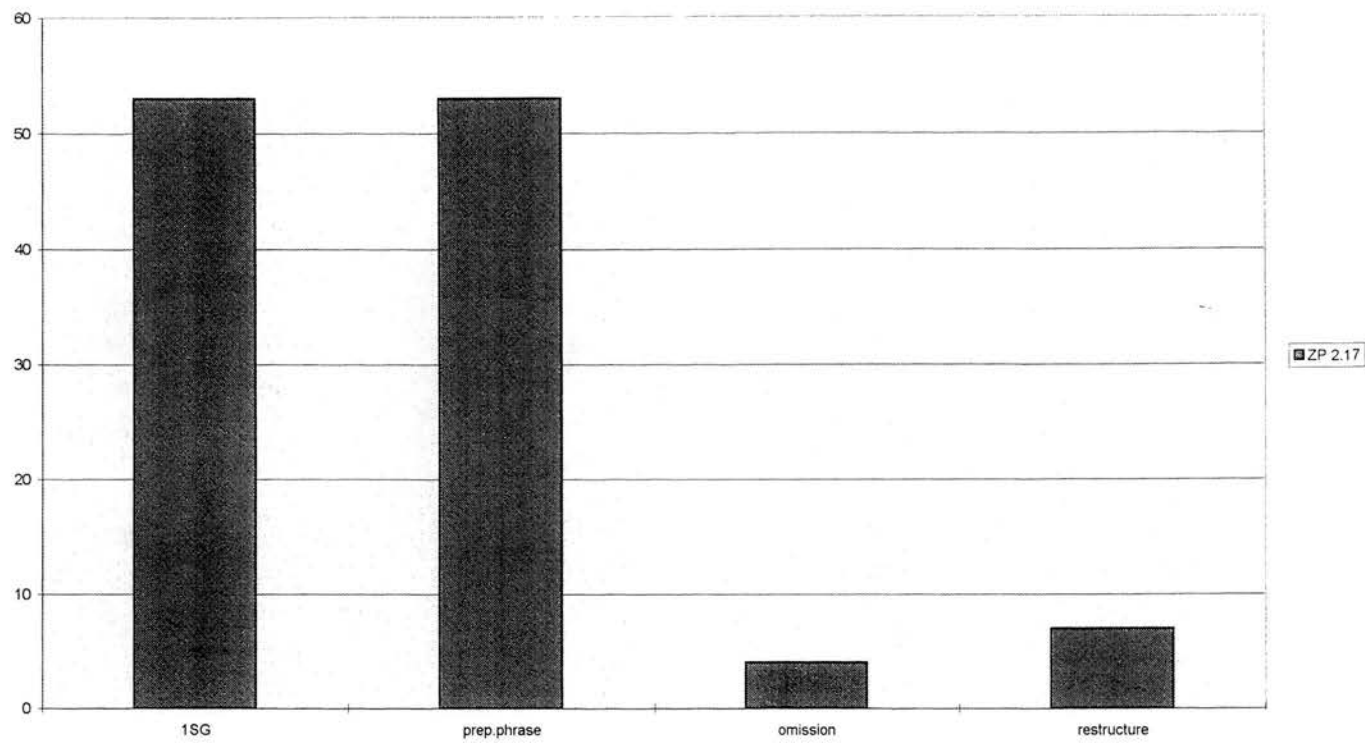
Interpretation of ZP 2.15



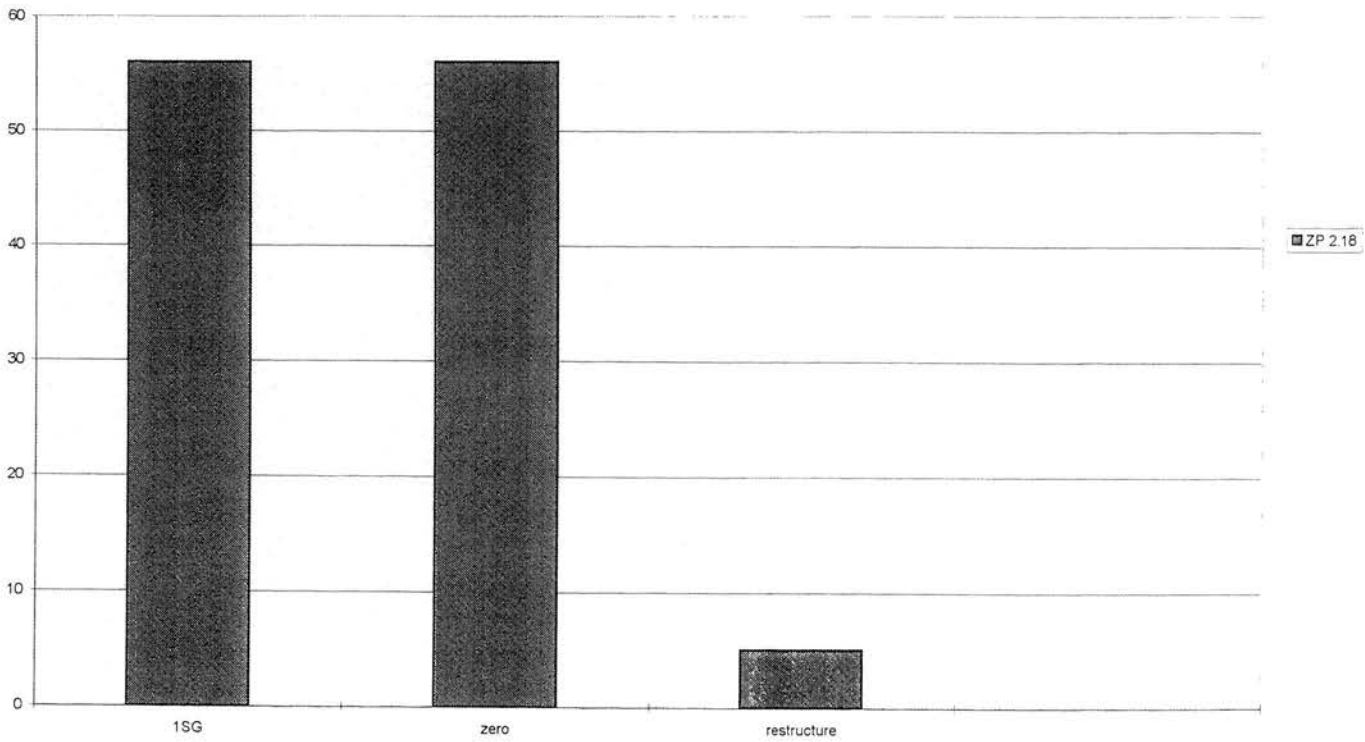
Interpretation of ZP 2.16



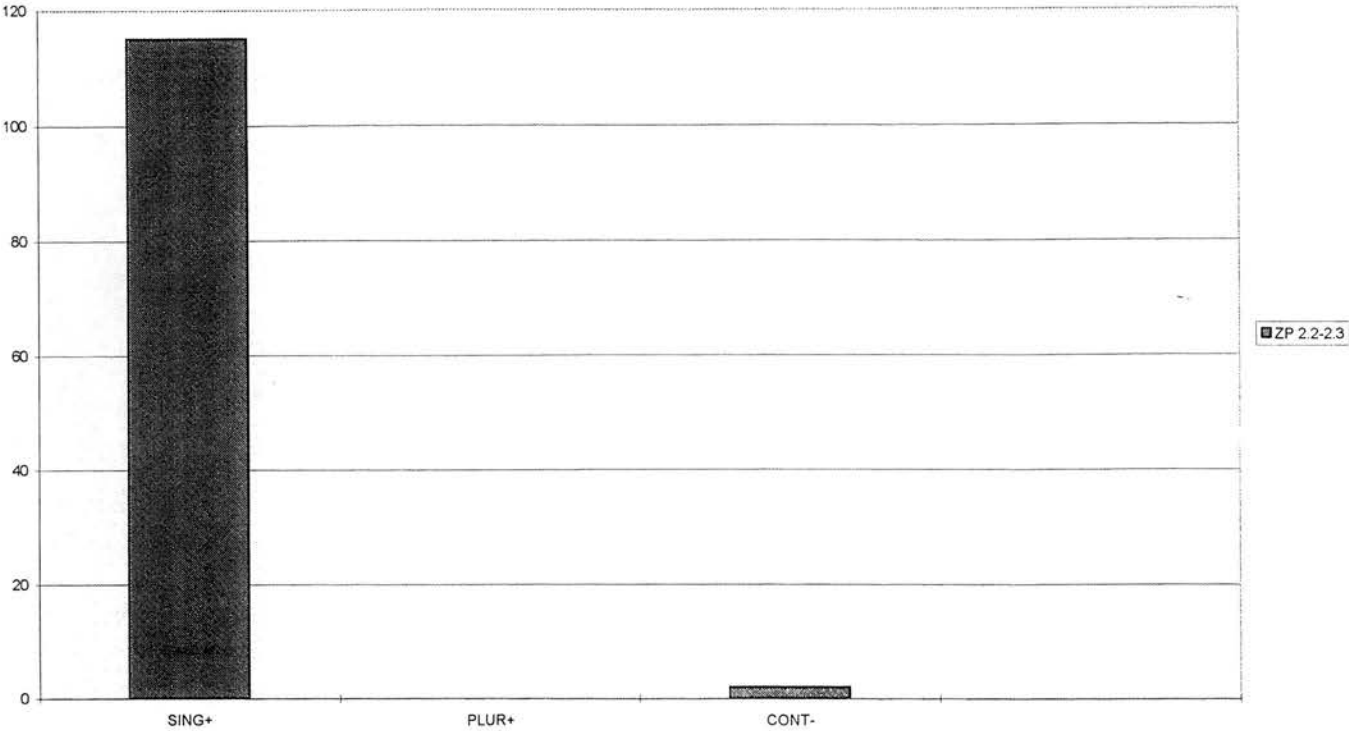
Interpretation of ZP 2.17



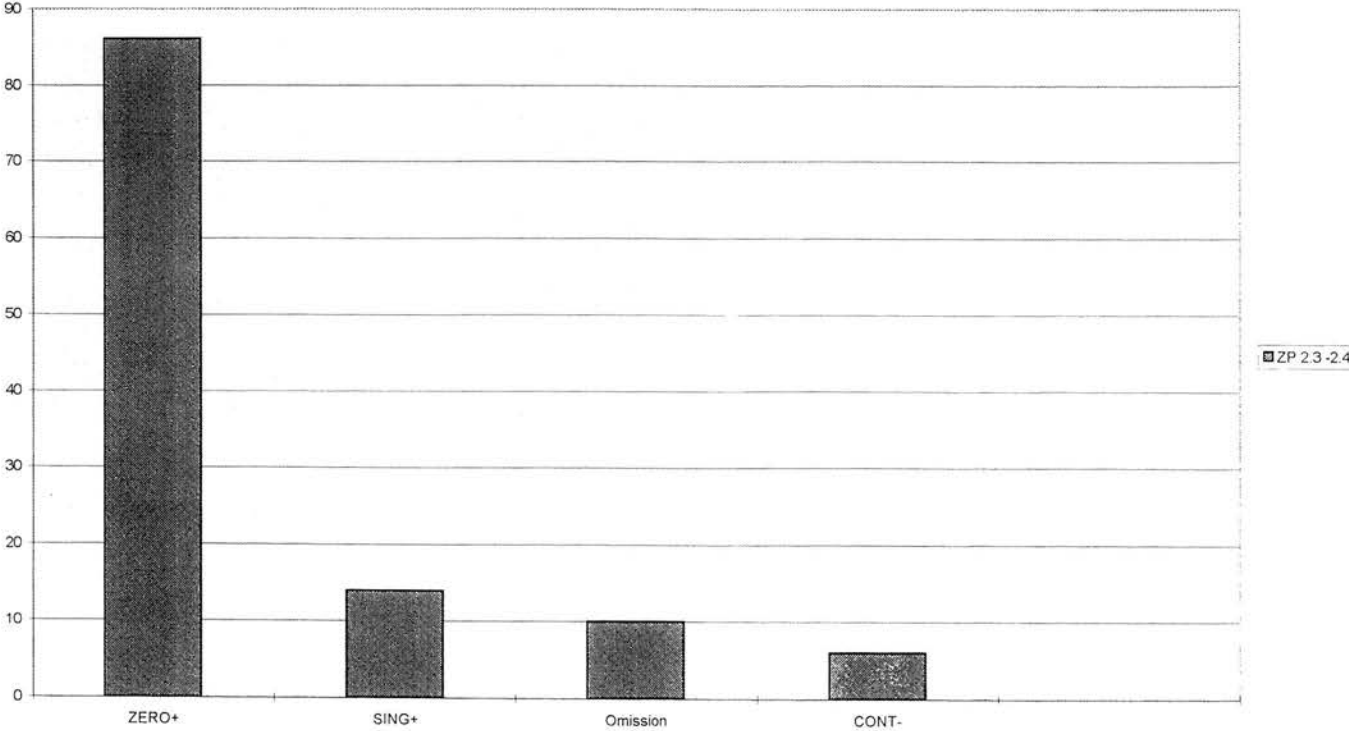
Interpretation of ZP 2.18



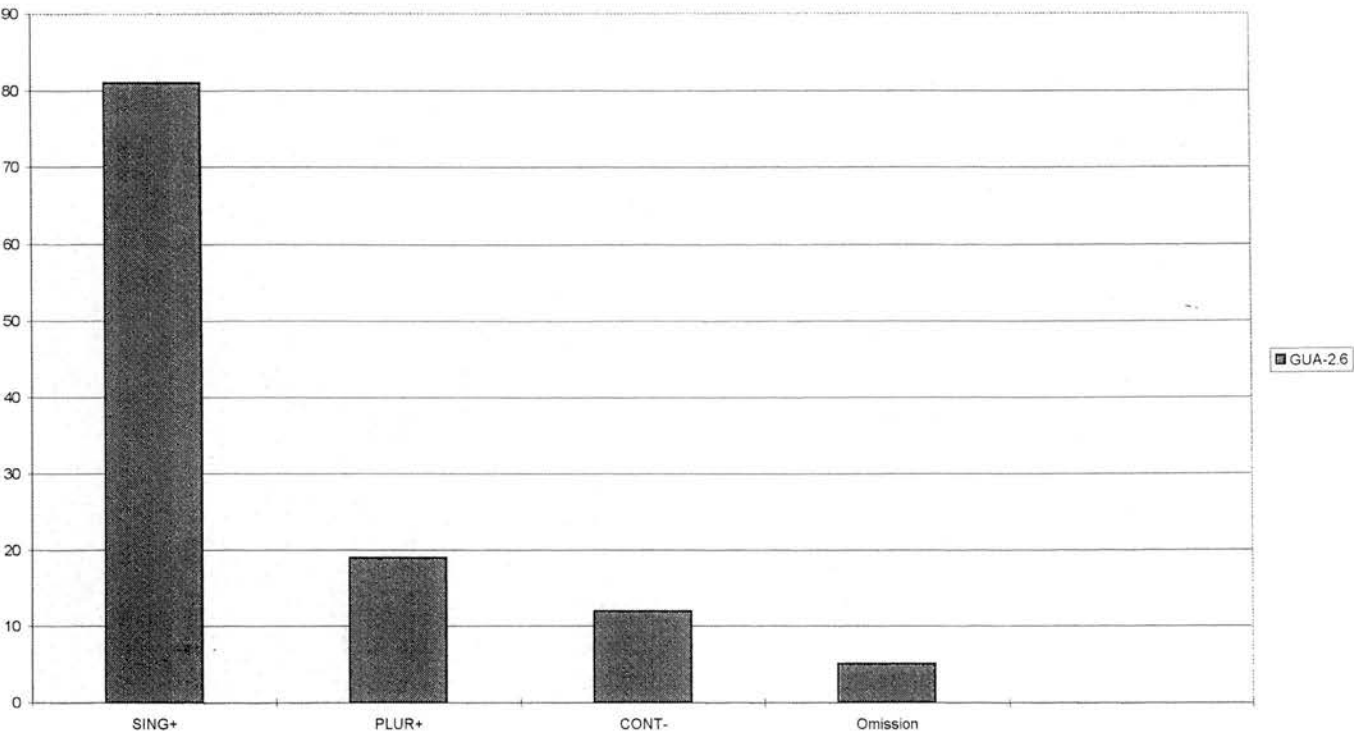
Continuity of reference ZP 2.2 - ZP 2.3



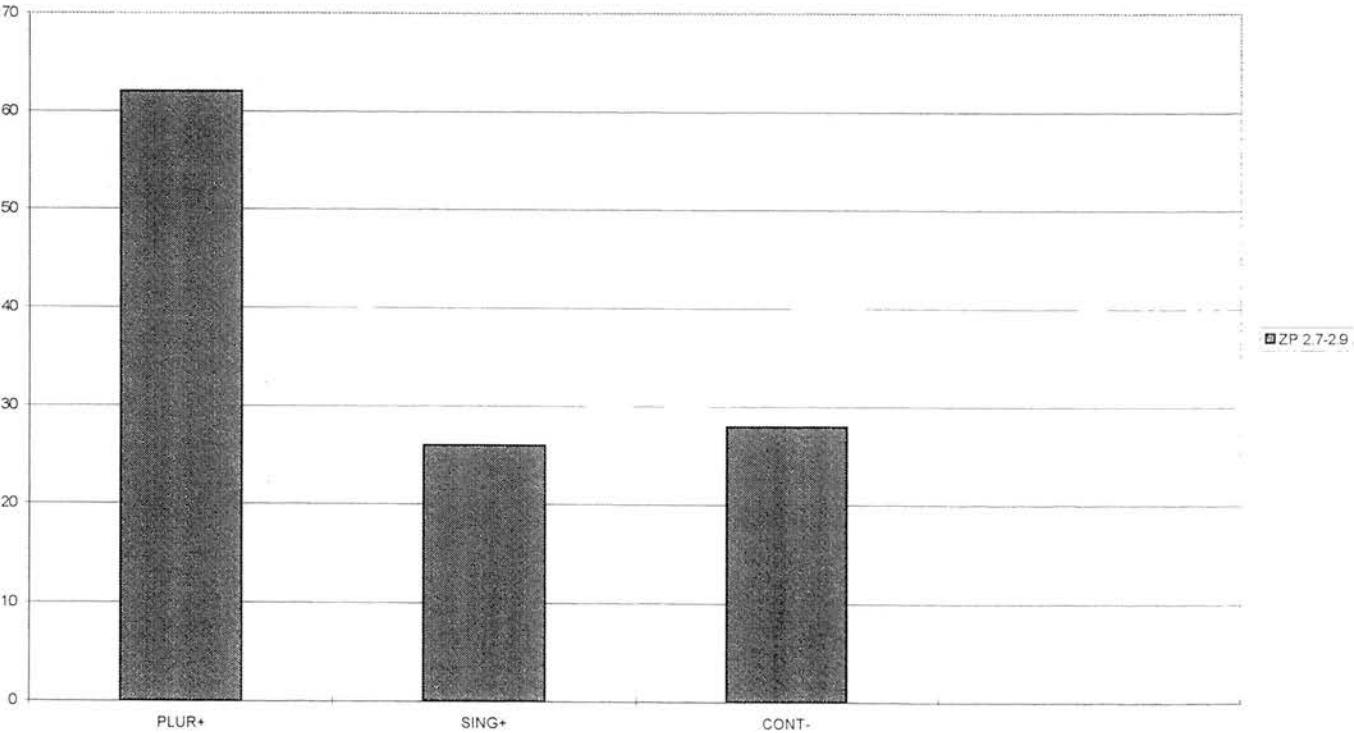
Continuity of reference ZP 2.3 - ZP 2.4



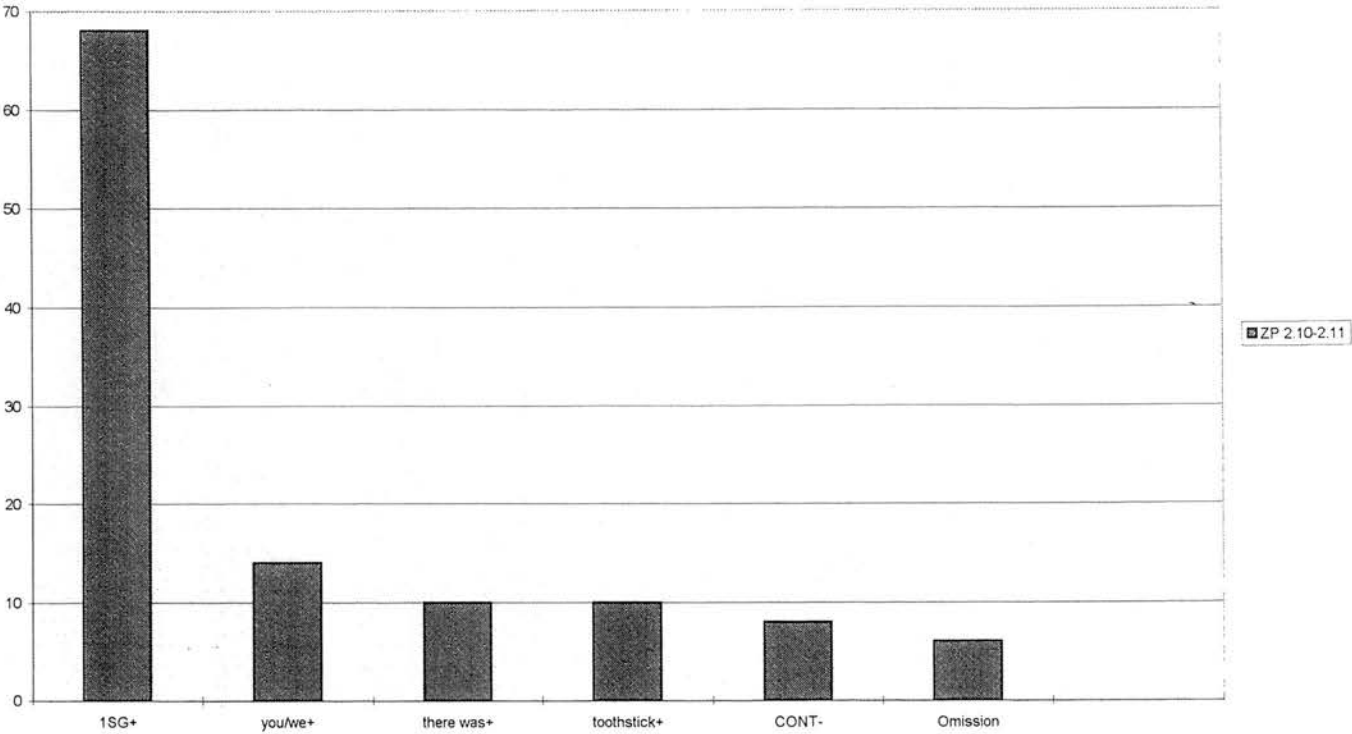
Continuity of reference 'guazi' - ZP 2.6



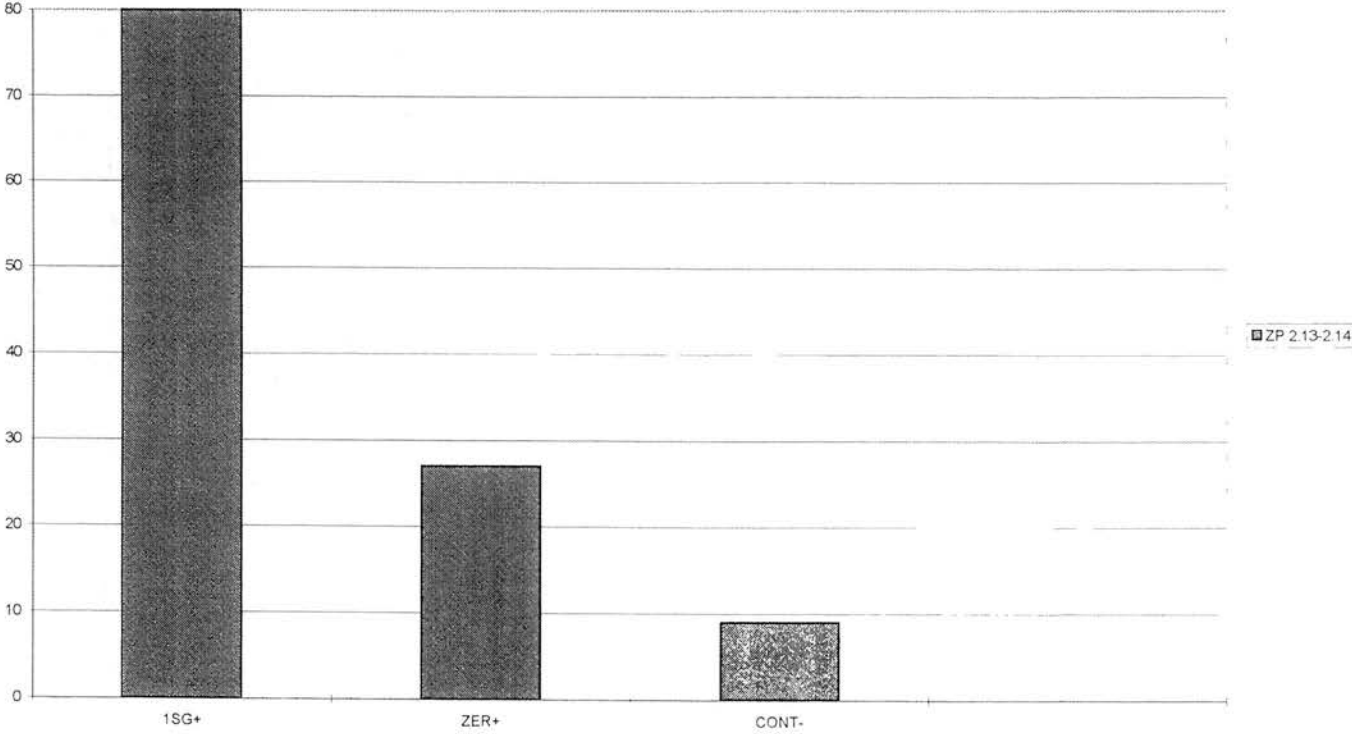
Continuity of reference ZP 2.7 - ZP 2.8 - ZP 2.9



Continuity of reference ZP 2.10 - ZP 2.11

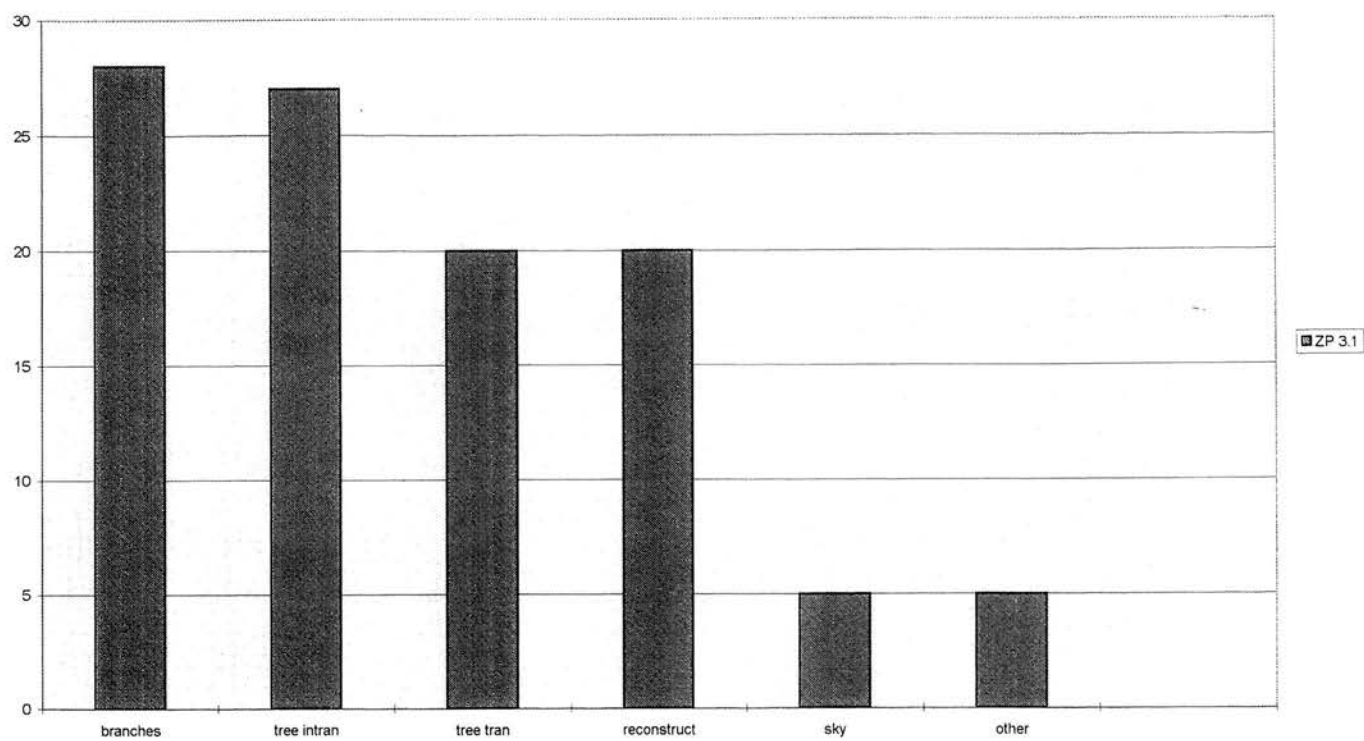


Continuity of reference ZP 2.13 - ZP 2.14

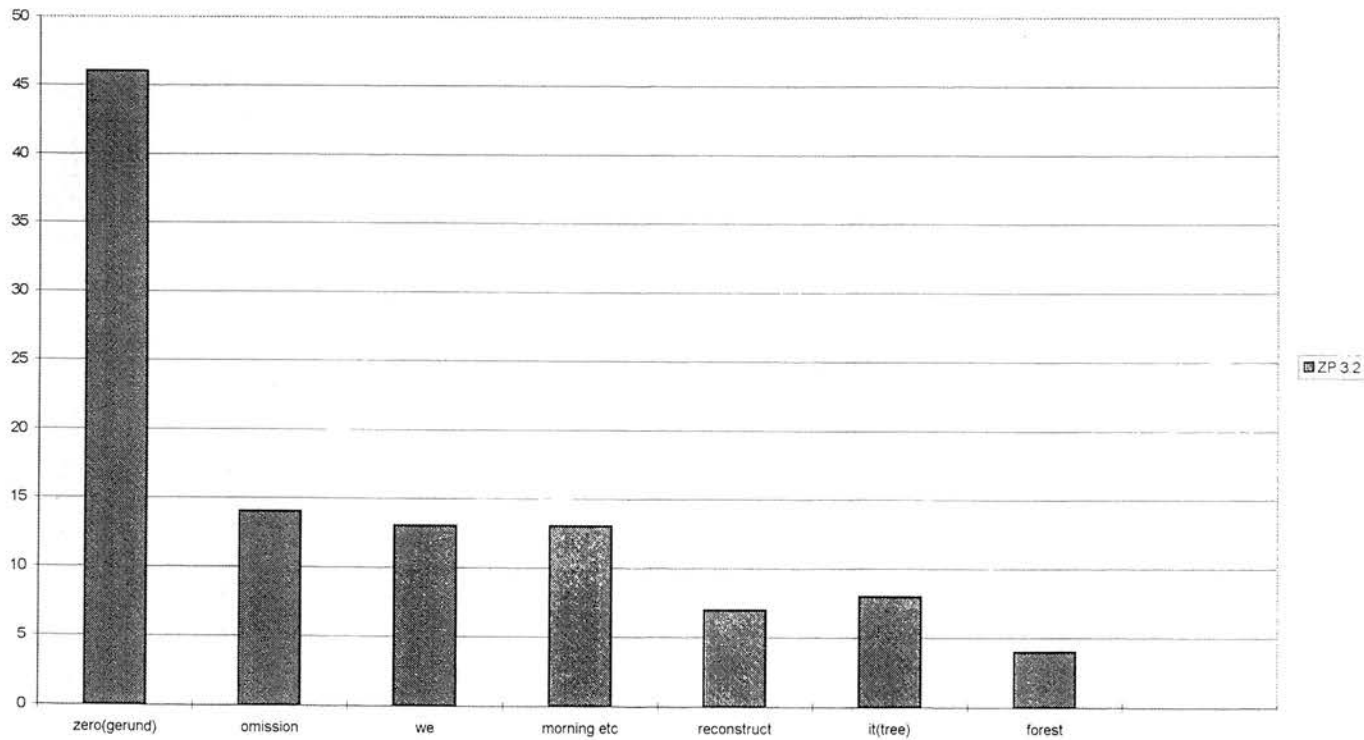




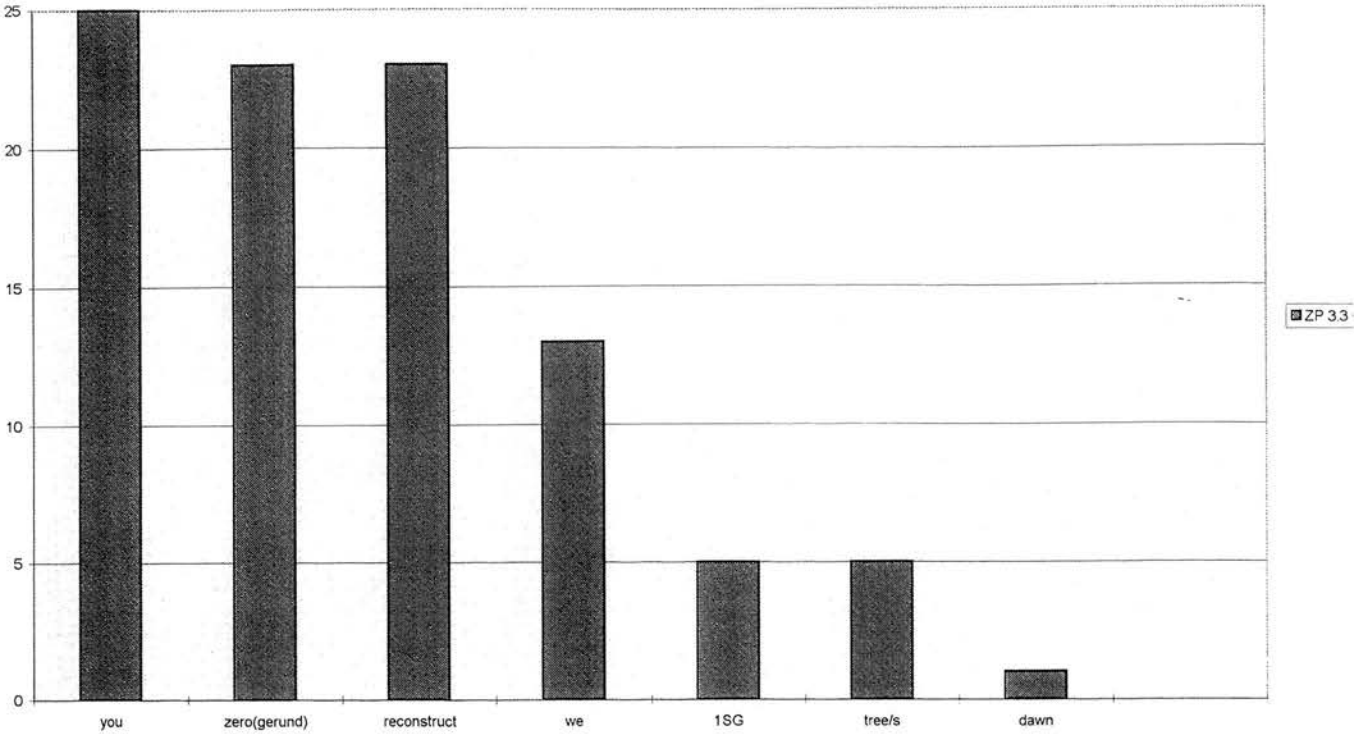
Interpretation of ZP 3.1



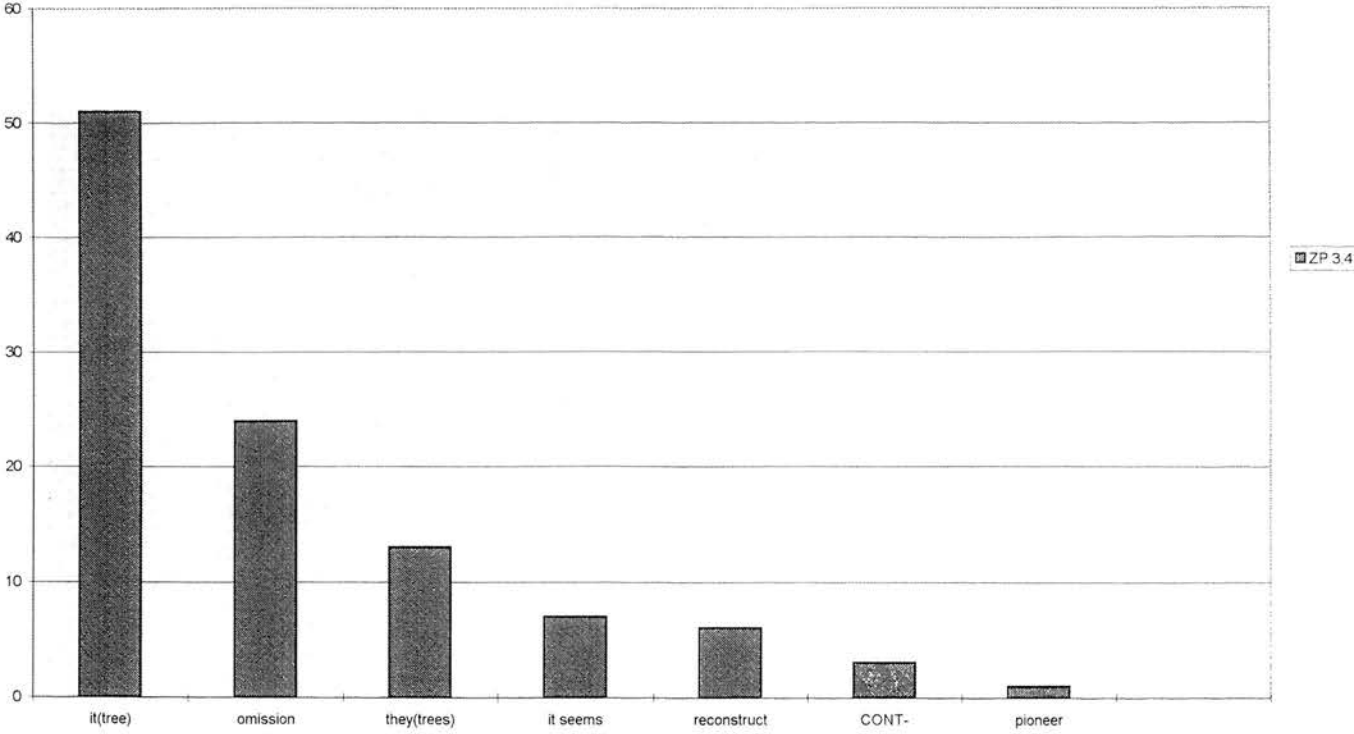
Interpretation of ZP 3.2



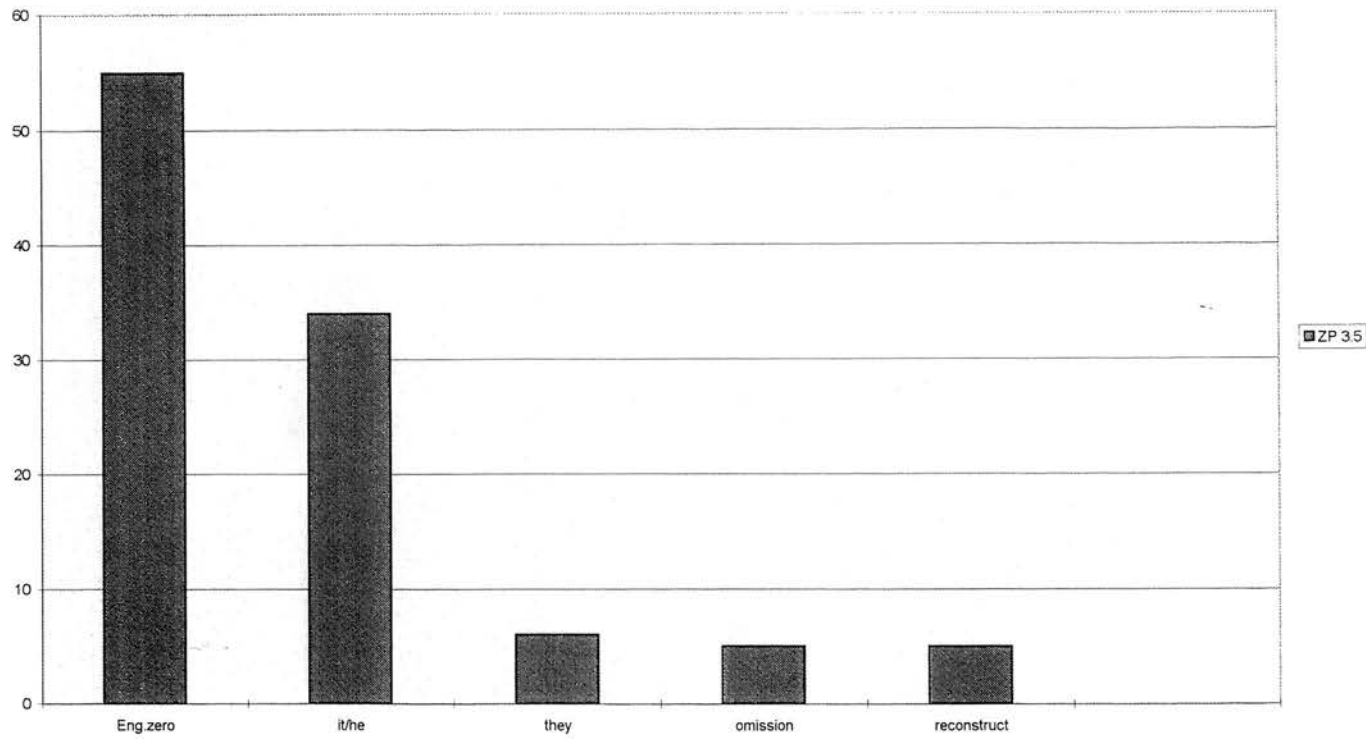
Interpretation of ZP 3.3



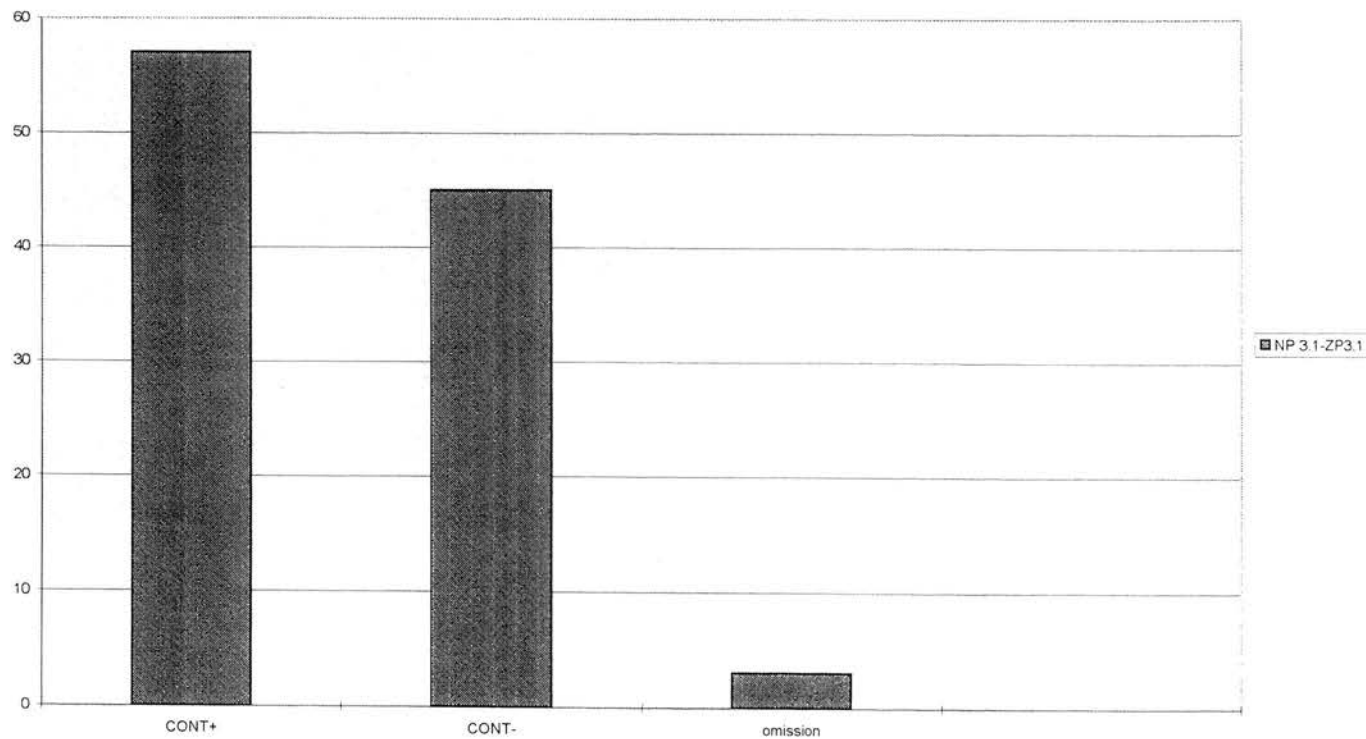
Interpretation of ZP 3.4



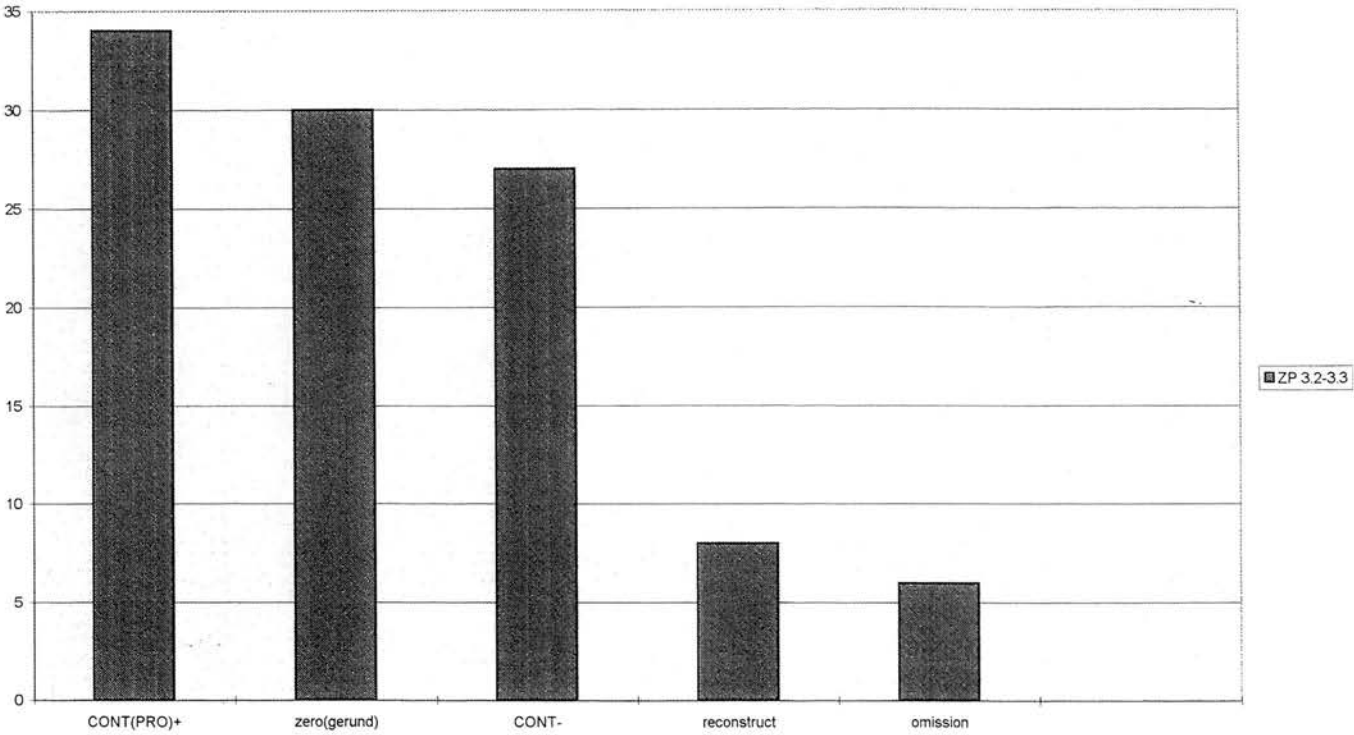
Interpretation of ZP 3.5



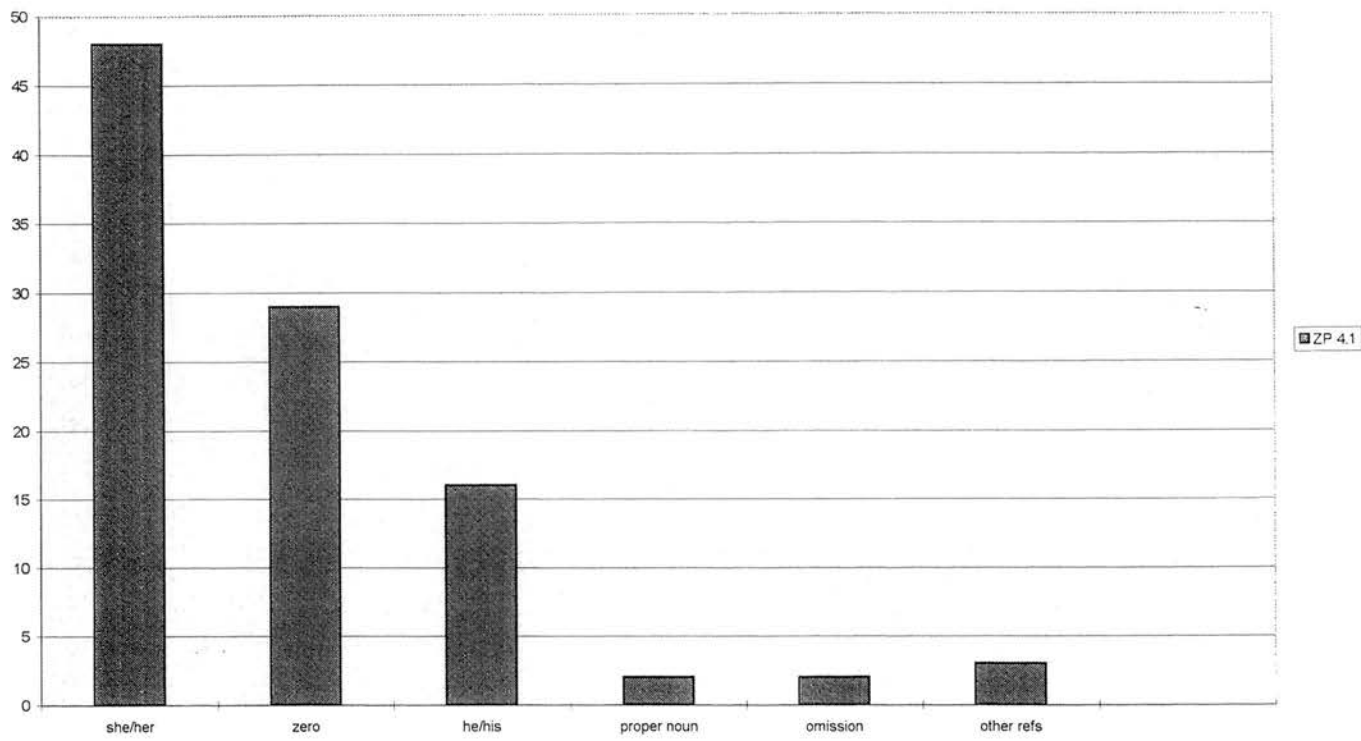
Continuity of reference NP 3.1 - ZP 3.1



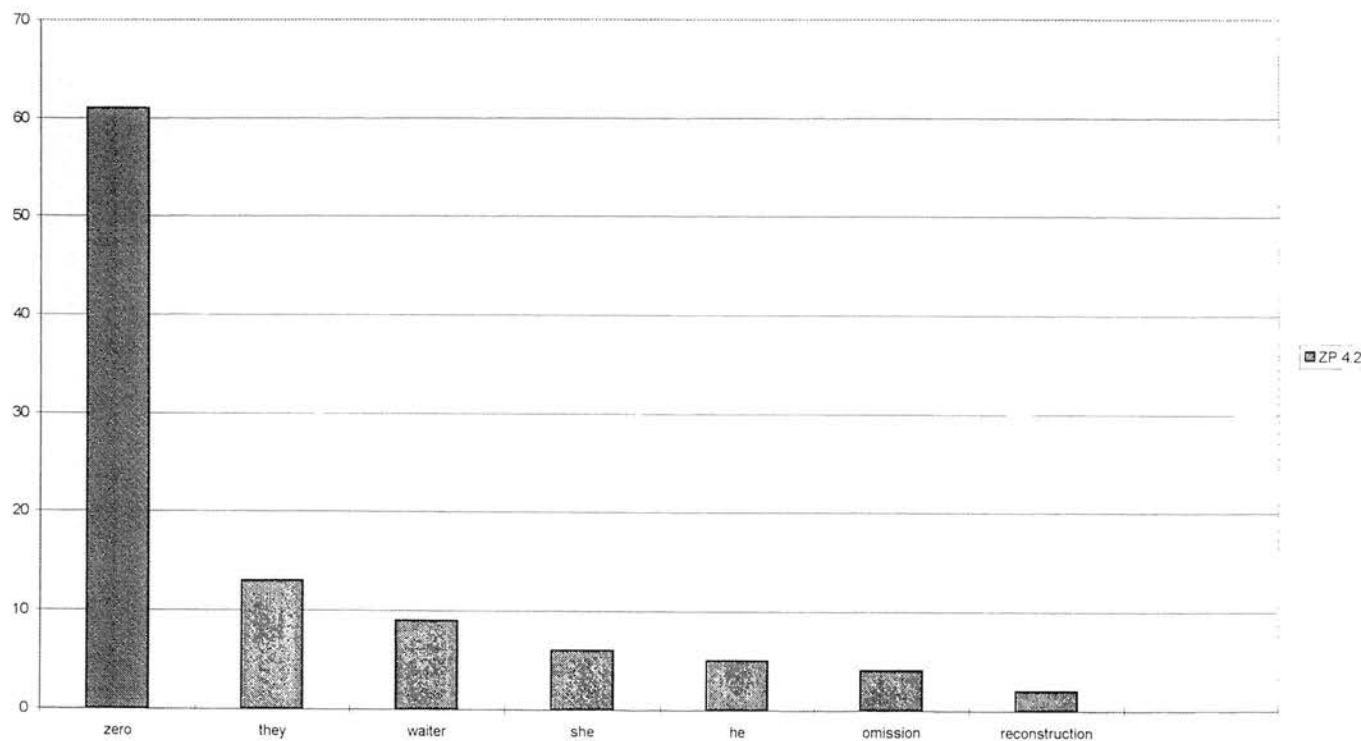
Continuity of reference ZP 3.2 - ZP 3.3



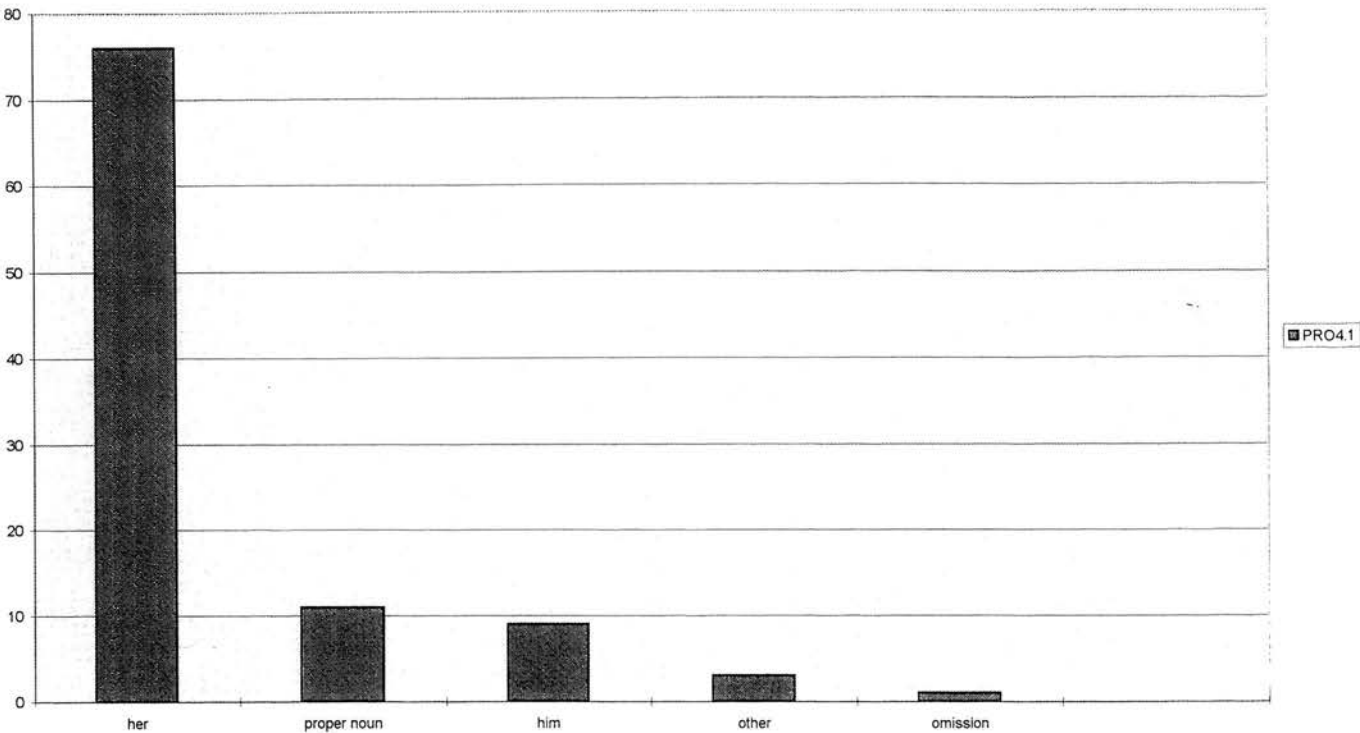
Interpretation of ZP 4.1



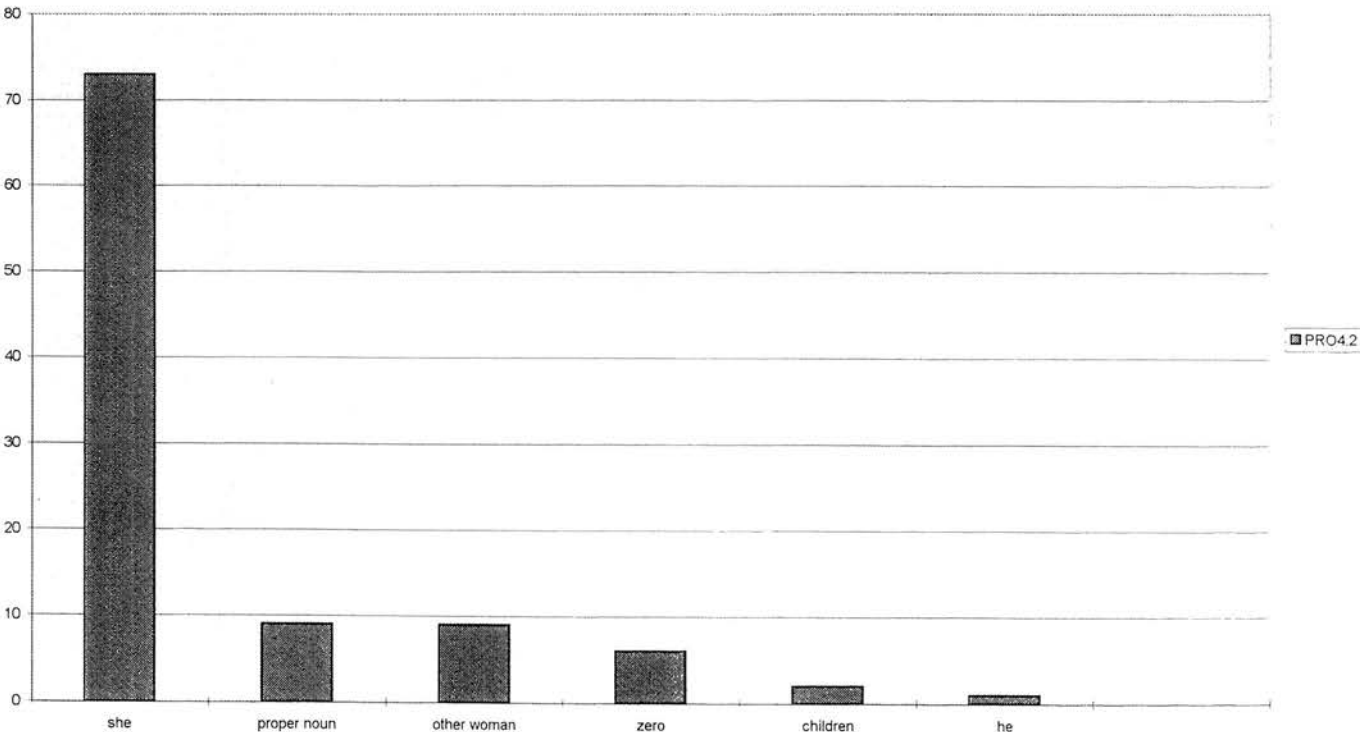
Interpretation of ZP 4.2



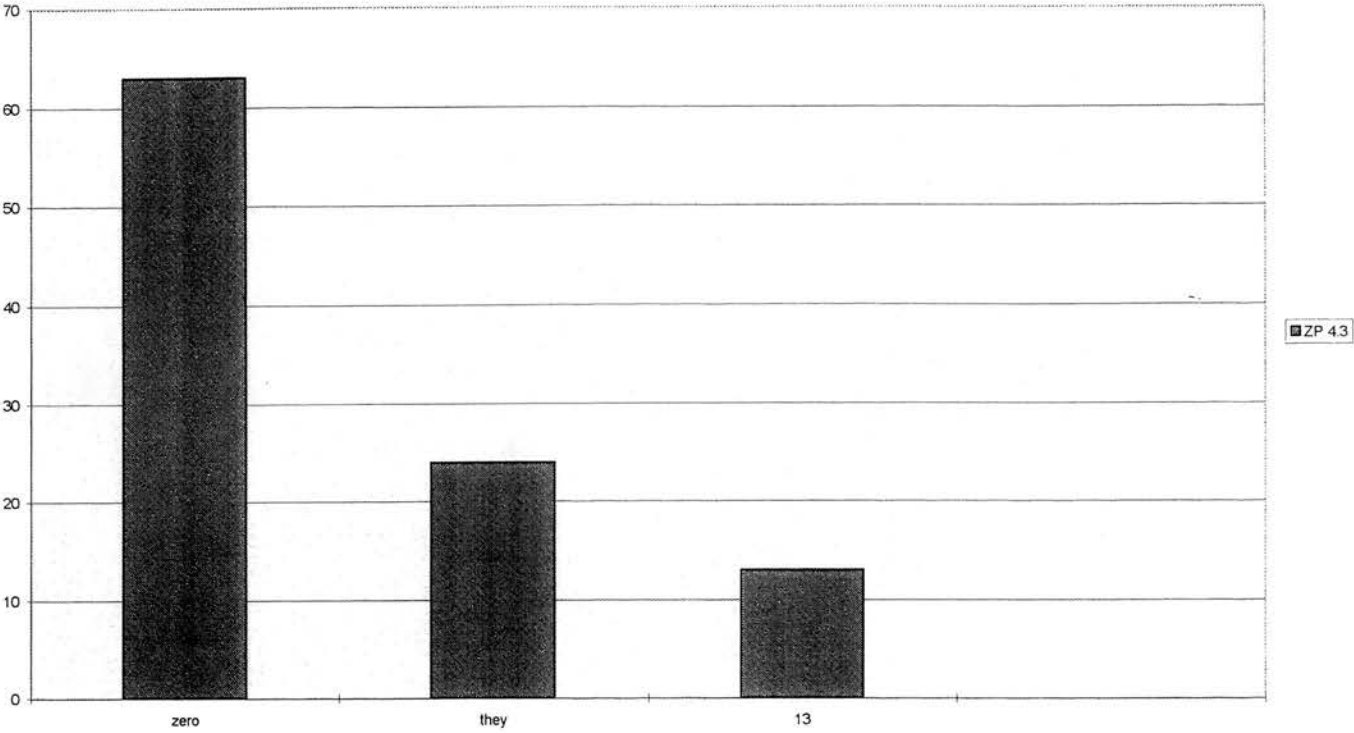
Interpretation of PRO 4.1



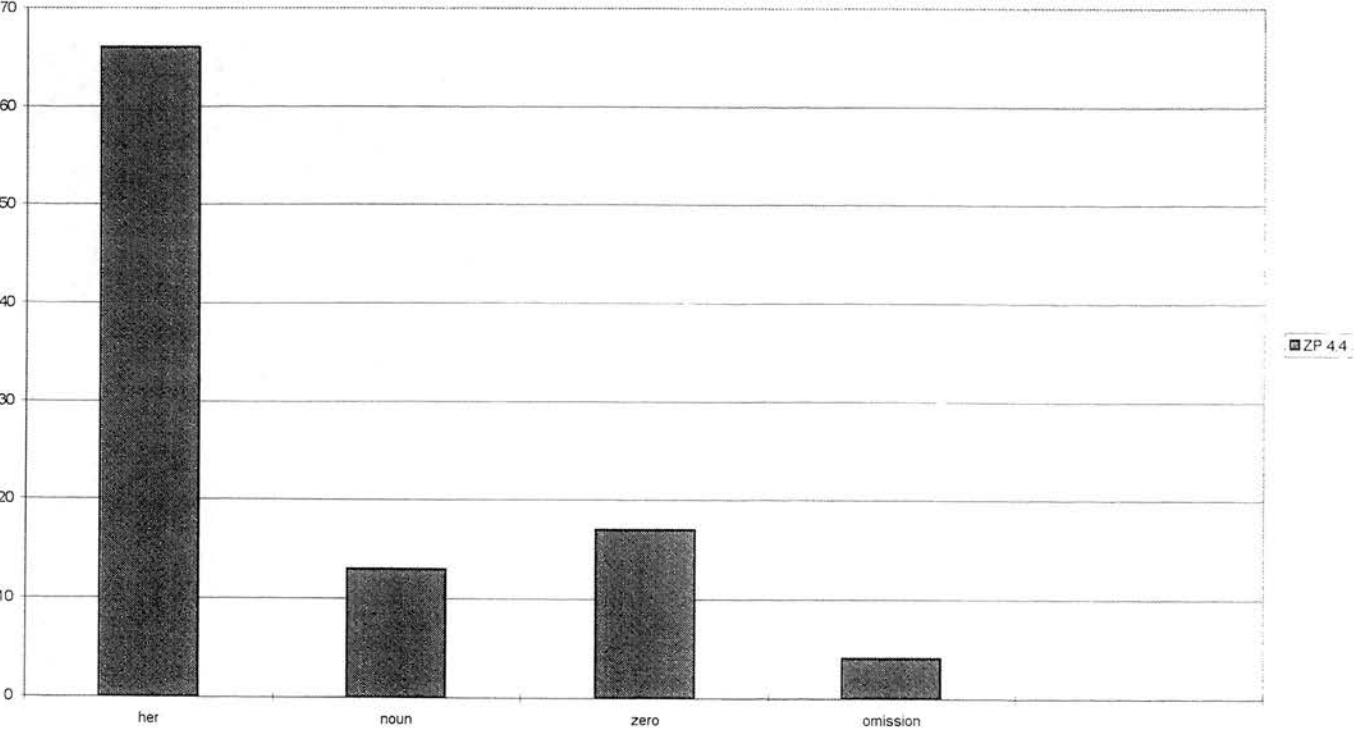
Interpretation of PRO 4.2



Interpretation of ZP 4.3

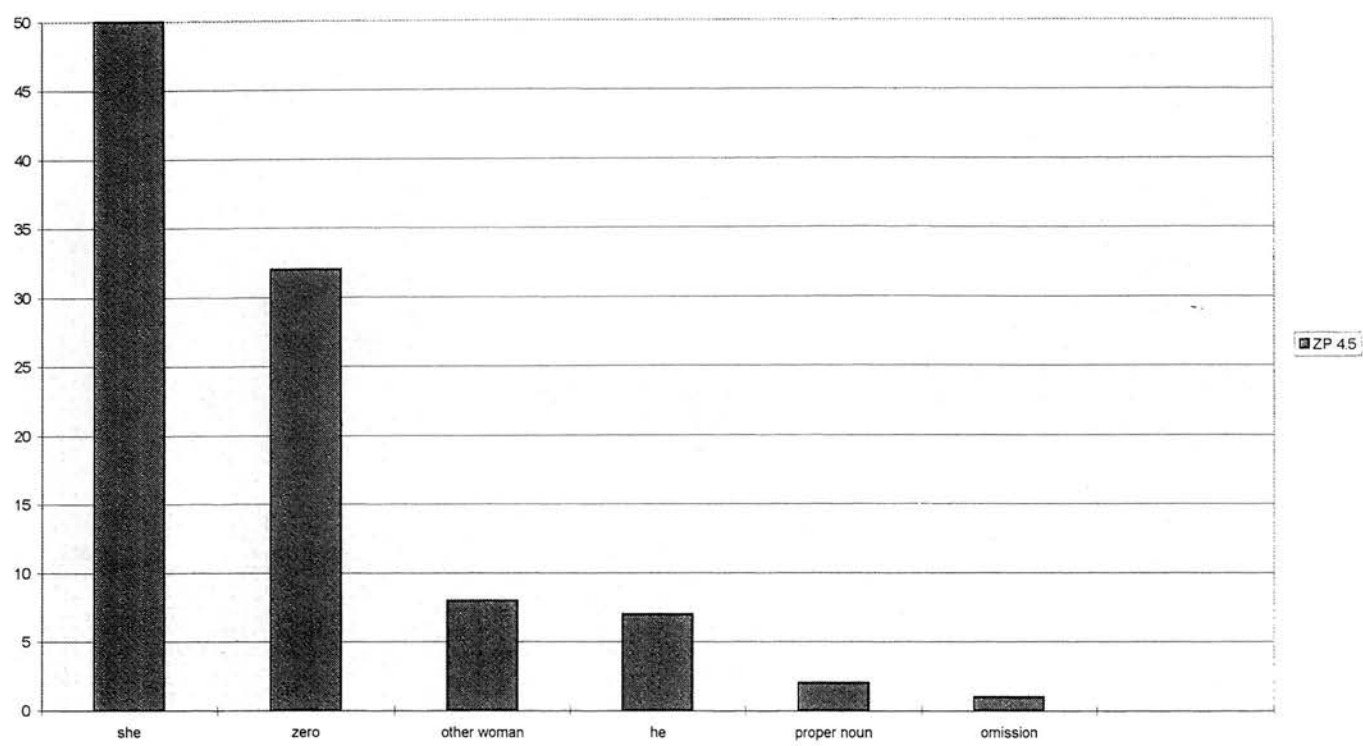


Interpretation of ZP 4.4

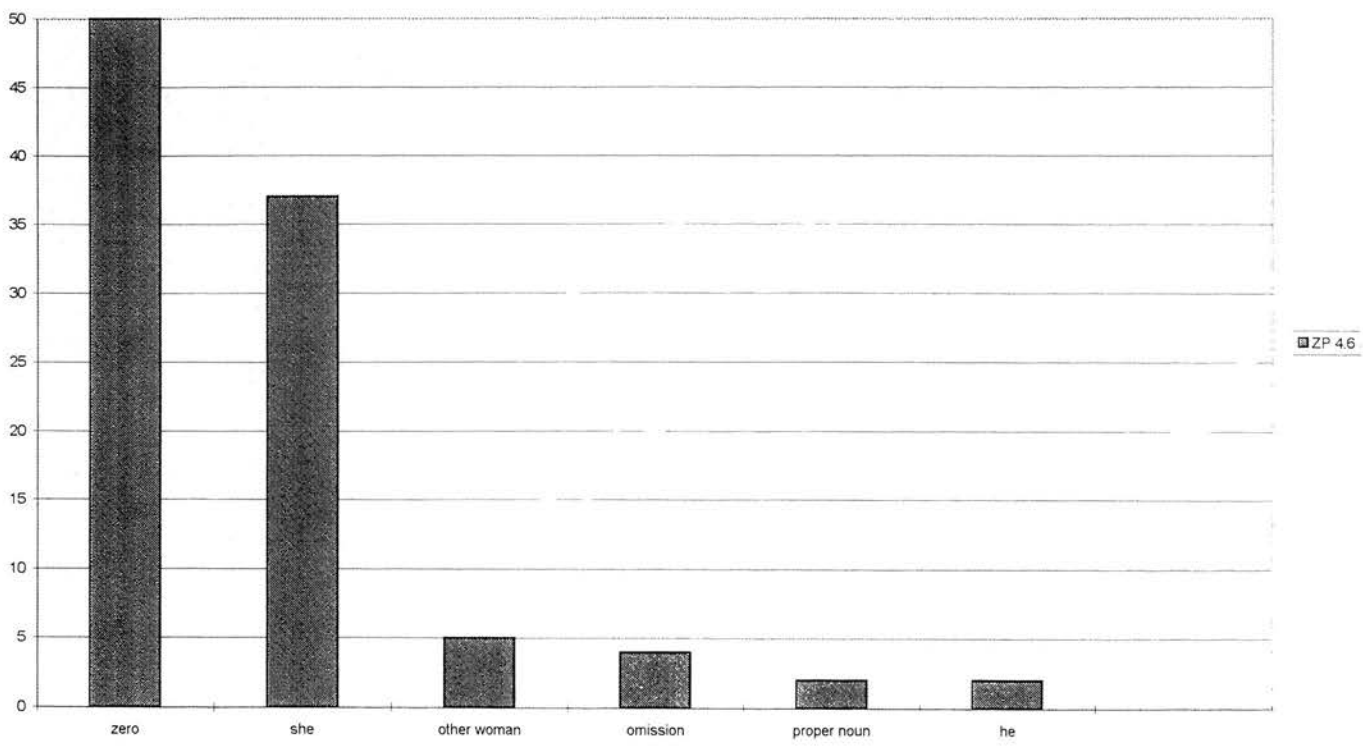




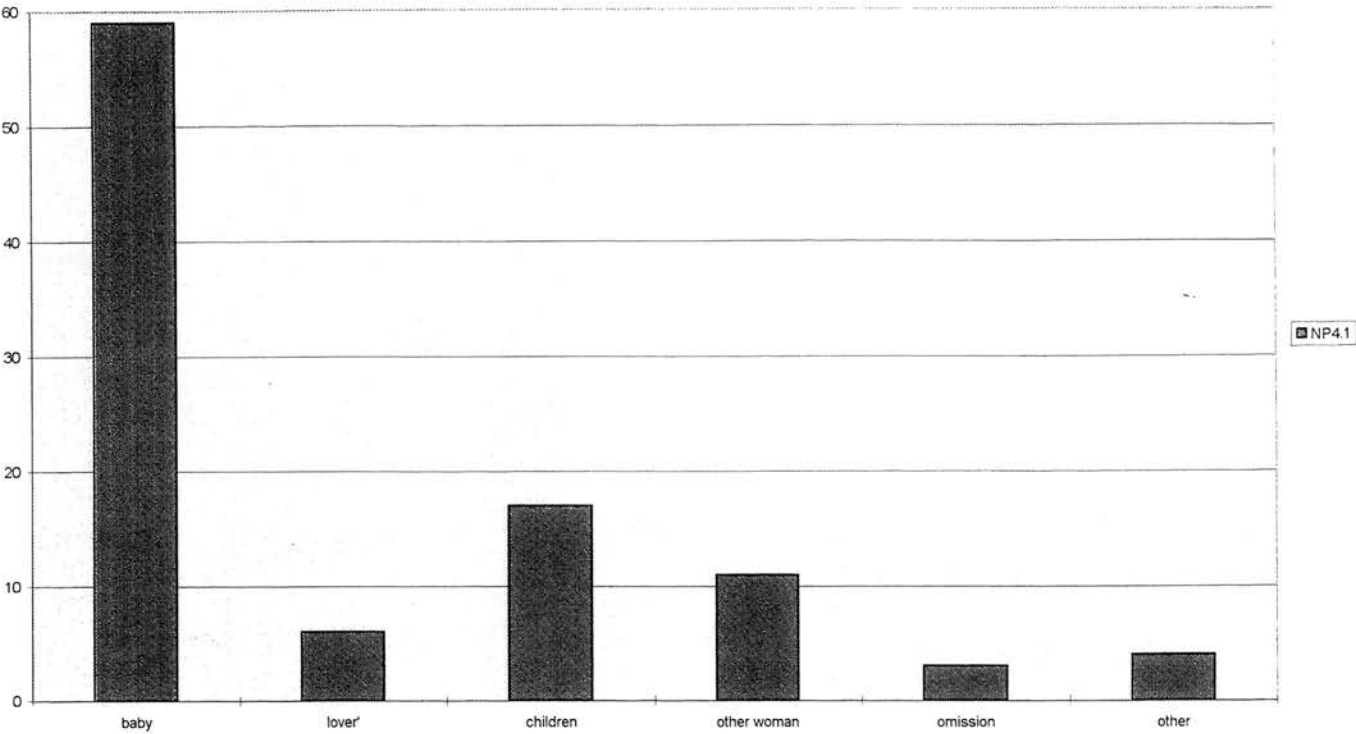
Interpretation of ZP 4.5



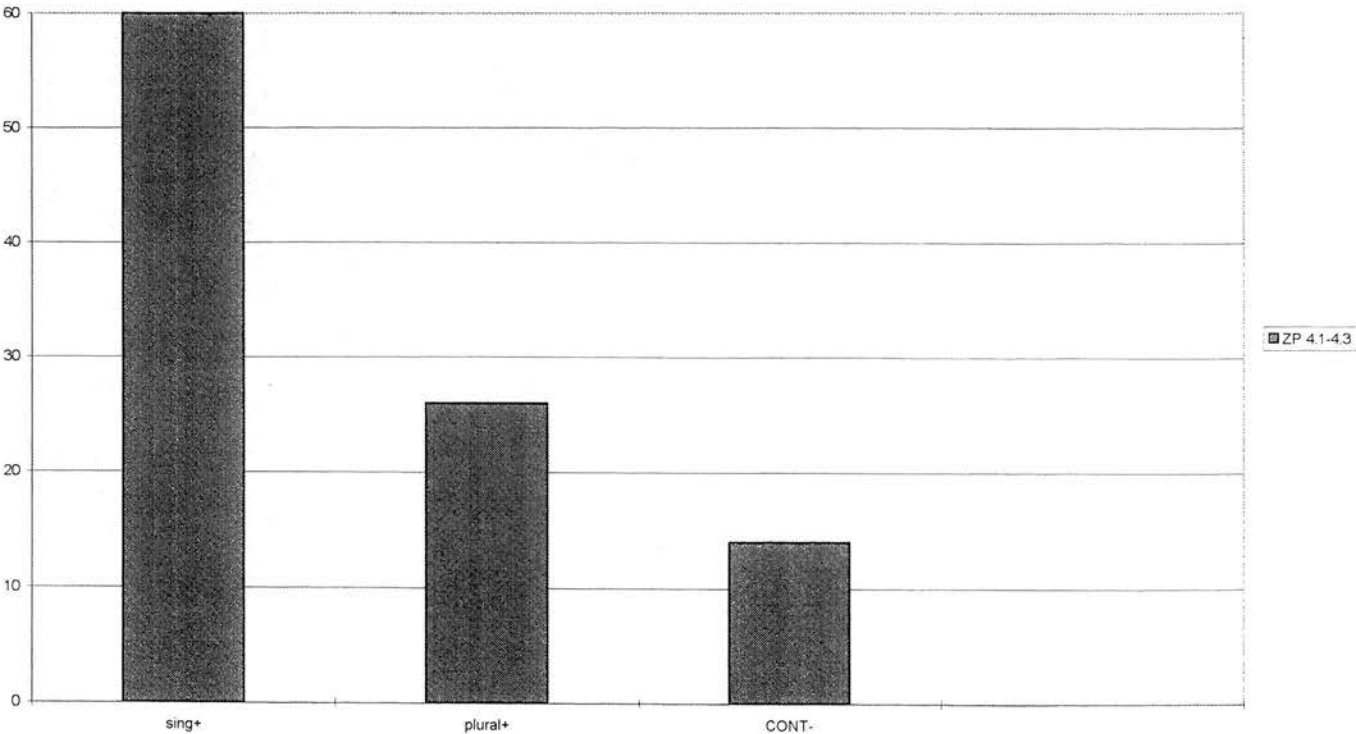
Interpretation of ZP 4.6



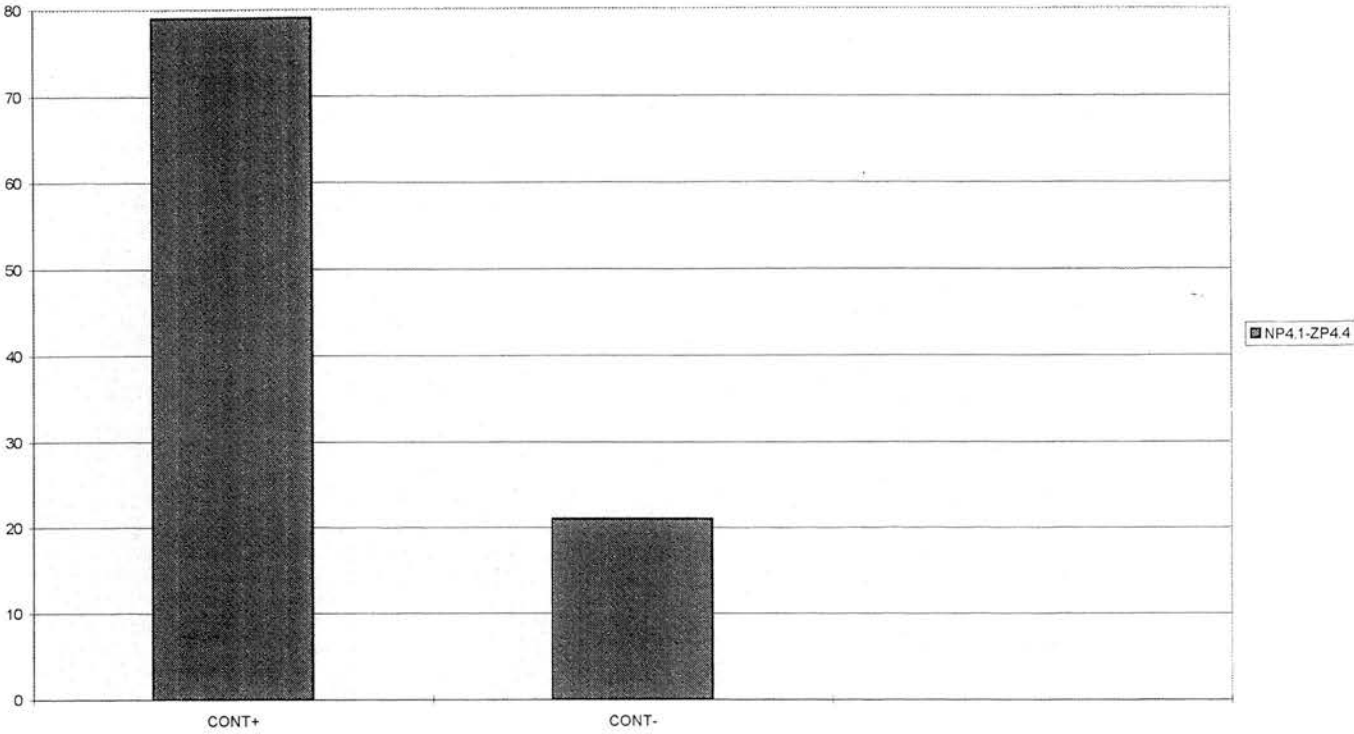
Interpretation of NP 4.1



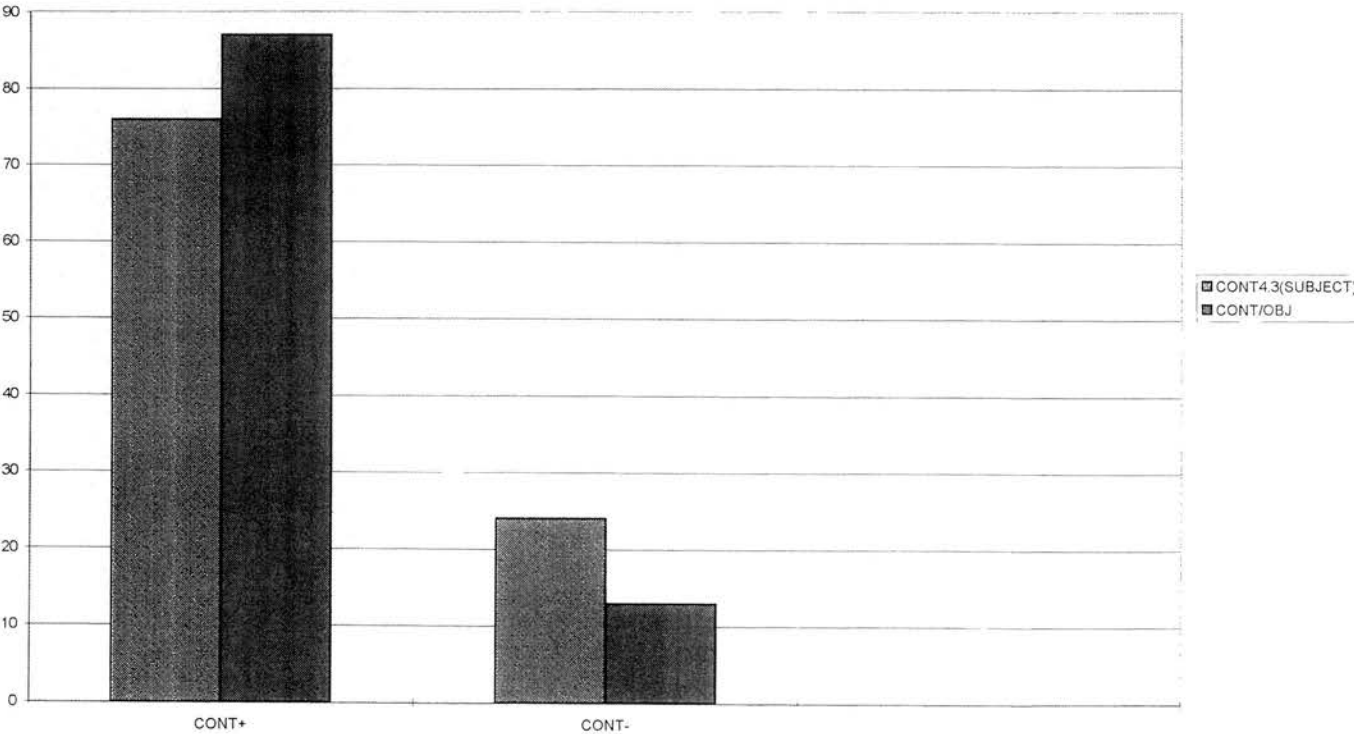
Continuity of reference ZP 4.1- ZP 4.2- ZP 4.3



Continuity of reference NP4.1- ZP 4.4



Continuity of reference NP 4.2- PRO 4.2- ZP 4.5-ZP 4.6



## **Appendix II: raters' questionnaires**

Note: no rater questionnaire was prepared for either Text 2 or Text 4, as it was felt on first inspection that subjects' anaphoric resolution in these texts was uniform and generally correct. It was only on later examination that variations were revealed. Analysis of solutions of Text 2 and Text 4 was therefore based on the researcher's intuitions derived from the ratings of Texts 1 and 3.

## Appendix 2: example of raters' questionnaire

### Accuracy judgement questionnaire: Text 1

#### 1. Please read the short text below:

一些摹擬魯迅先生雜文的人，往往忘卻了這些雜文產生的時間和空間，目的和對象。在他的雜文中，有些地方不得已的“咬文嚼字”，“拐彎抹角”，“引經據典”，“聲東擊西”，“文白混用”，或故意“含糊其辭”，都是當時的年代、環境的產物，連先生自己也認為詬病的。如今我們不能盲目地效仿，應該只是在需要的時候才能夠用一用它。我們所主要該學的，卻是那及時磨練武器、精通武藝、以及臨敵無畏的精神。

#### 2. Please read the following translations of phrases from the text above and rate them on a scale of 5 *for accuracy in the context of the text as a whole* and insert number 1,2,3,4, or 5 in the boxes on the right.

- 5 = accurate;
- 4 = fairly close;
- 3 = not entirely wrong but could be closer;
- 2 = inaccurate;
- 1 = inaccurate and irrelevant.

Text 1, Phrase A:

在他的雜文中，有些地方不得已的

- a. In his essays there were some places where he could not help..... ☐
- b. In his essays there is something unavoidable like..... ☐
- c. In Mr Low's writings there were some features like..... ☐
- d. In his essays there are many places we cannot use..... ☐
- e. In some of his works there are inevitable..... ☐
- f. His articles tend to be ..... ☐
- g. Some dragons were used in that time. ☐

Comments/introspection on translation of Phrase A:

Text 1, Phrase B:

連先生自己也認為

- |    |  |                          |
|----|--|--------------------------|
| a. | Even Mr. Lu Xun thought....  | <input type="checkbox"/> |
| b. | The teacher himself thought that.....  | <input type="checkbox"/> |
| c. | Even teachers themselves thought this.....                                   | <input type="checkbox"/> |
| d. | Even those teachers at that time thought.....                                | <input type="checkbox"/> |
| e. | Even our teacher thinks.....   | <input type="checkbox"/> |
| f. | They (the people who copy Lu Xun) thought it was some problem of themselves. | <input type="checkbox"/> |
| g. | If we can spot.....  | <input type="checkbox"/> |
| h. | Even myself also think it is .....   | <input type="checkbox"/> |
| i. | The scriptwriters themselves also thought.....                               | <input type="checkbox"/> |

Comments/introspection on translation of Phrase B:



Text 1, Phrase C.

認為詬病的

- |    |   |                          |
|----|---|--------------------------|
| a. | ....he reckoned these were flaws.                                     | <input type="checkbox"/> |
| b. | ....he felt that is a mistake.  | <input type="checkbox"/> |
| c. | ....he felt that it is not the proper grammar.                        | <input type="checkbox"/> |
| d. | ....the school teacher wouldn't understand his mistakes.              | <input type="checkbox"/> |
| e. | ....Mr. Lu Xun himself believes the people who copied him were wrong. | <input type="checkbox"/> |
| f. | .....he himself was sick.   | <input type="checkbox"/> |
| g. | ...he was ill about them.   | <input type="checkbox"/> |
| h. | ..if we can spot where we go wrong....                                | <input type="checkbox"/> |
| i. | ....he finds that his magazine had a lot of problems.                 | <input type="checkbox"/> |

Comments/introspection on translation of Phrase C.

Text 1, Phrase D.

如今我们不能盲目地效仿，

- |    |  |                          |
|----|--|--------------------------|
| a. | We cannot just copy it blindly.                            | <input type="checkbox"/> |
| b. | We can't copy his work without consideration.              | <input type="checkbox"/> |
| c. | We should not use them without understanding them.         | <input type="checkbox"/> |
| d. | We cannot copy things blindly.                             | <input type="checkbox"/> |
| e. | We cannot totally follow the style.                        | <input type="checkbox"/> |
| f. | We shouldn't try following their style.                    | <input type="checkbox"/> |
| g. | We cannot follow blindly.                                  | <input type="checkbox"/> |
| h. | We shall not follow the wrong path of the past.            | <input type="checkbox"/> |
| i. | We can't do without eyes and brain.                        | <input type="checkbox"/> |
| j. | We don't need to learn from his magazine without thinking. | <input type="checkbox"/> |
| k. | We can't really learn kung-fu for the sake of it.          | <input type="checkbox"/> |

Comments/introspection on translation of Phrase D:

Text 1, Phrase E:

我们所主要学习的，

- |    |   |                          |
|----|---|--------------------------|
| a. | The main thing that we have to learn is...          | <input type="checkbox"/> |
| b. | What we should learn....                            | <input type="checkbox"/> |
| c. | The main things we have to learn are....            | <input type="checkbox"/> |
| d. | The main purpose of what we have to learn is...     | <input type="checkbox"/> |
| e. | The main theme we need to learn from his work is... | <input type="checkbox"/> |
| f. | The most important principle for us to learn is.... | <input type="checkbox"/> |
| g. | The major subject we have to learn is....           | <input type="checkbox"/> |
| h. | All we need to learn is.....                        | <input type="checkbox"/> |
| j. | We have one thing to learn, that is....             | <input type="checkbox"/> |
| k. | We should learn what is needed.....                 | <input type="checkbox"/> |

Comments/introspection on translation of Phrase E.

Accuracy judgement questionnaire: Text 3

Please read the short text below:

当幼树从松林的废墟上破土而出，在天边舞动那皎洁的苗条，这不象是又回到了破晓的黎明，看到了新一代的大森林露出了它灿烂夺目的曙光吗！凭着它那纤弱苍白的一束嫩条，幼树是多么坚毅和勇敢的先驱者，似乎是不自量力竟要在一片荒漠贫瘠的土地上从头拓荒，要重建一座比已经有过的都更加伟大的森林。它又是怎样达观和睿智的开拓者，还要荫庇和哺育那按照大自然的安排将要取代它的松林。

1. Please read the following translations of phrases from the text above and rate them on a scale of 5 *for accuracy in the context of the text as a whole* and insert numbers 1,2,3,4,or 5 in the boxes on the right.

5 = accurate;

4 = fairly close;

3 = not entirely wrong but could be closer;

2 = inaccurate;

1 = inaccurate and irrelevant.

Text 3, Phrase A:

当幼树从松林的废墟上破土而出，在天边舞动那皎洁的苗条，

- |    |   |                          |
|----|---|--------------------------|
| a. | .. the young tree is waving its branch in the sky..     | <input type="checkbox"/> |
| b. | ...the small tree...is dancing in the sky...            | <input type="checkbox"/> |
| c. | ..... the thin stem sways across the sky..              | <input type="checkbox"/> |
| d. | ..the young tree dances with the seedlings in the sky.. | <input type="checkbox"/> |
| e. | ..the grass was dancing in the sky..                    | <input type="checkbox"/> |
| f. | ..the sky is dancing smoothly..                         | <input type="checkbox"/> |
| g. | ..the wind blows the small tree..                       | <input type="checkbox"/> |
| h. | ...the sky is white in colour..                         | <input type="checkbox"/> |
| i. | ...the sun was coming, dancing high over the sky..      | <input type="checkbox"/> |
| j. | ..you see the little branches..                         | <input type="checkbox"/> |

Comments/introspection on translation of Phrase 3.A

Text 3, Phrase B:

这不象是又回到了破晓的黎明，

- |    |  |                          |
|----|--|--------------------------|
| a. | ..isn't this like going back to the sunrise..          | <input type="checkbox"/> |
| b. | ...it is like the return to the sunrise..              | <input type="checkbox"/> |
| c. | ...like the sunrise is coming back..                   | <input type="checkbox"/> |
| d. | ...it is (like) the sunrise..                          | <input type="checkbox"/> |
| e. | ...when the sun breaks the sky at sunrise..            | <input type="checkbox"/> |
| f. | ..it (the young tree) is going back to the sunrise.... | <input type="checkbox"/> |
| g. | ..we have come back to the sunrise...                  | <input type="checkbox"/> |
| h. | ..like the small tree brings the sunrise back..        | <input type="checkbox"/> |
| i. | ..it is like everything is finished...                 | <input type="checkbox"/> |
| j. | ..the new moon changed to the sun ...                  | <input type="checkbox"/> |

Comments/introspection on translation of Phrase 3.B.

Text 3, Phrase C:

看到了新一代的大森林

- a. ..and seeing the new generation of forest.... ☐
- b. ...we can see the new generation of forest.... ☐
- c. Can you see the new generation of forest....? ☐
- d. You can see the new generation of forest... ☐
- e. Look at the new generation of forest...! ☐
- f. ..I see the new generation of forest.. ☐
- g. ..it (the young tree) can see the new generation of forest... ☐
- h. Does the young tree show the new generation of forest...? ☐
- I. Can the new generation of the big forest be seen..? ☐

Comments/introspection on translation of Phrase 3. C:



Text 3, Phrase D

似乎是不自量力

- a. It seems like the young tree doesn't consider its own strength... ☐
- b. It seems to have overestimated its own abilities... ☐
- c. ..as if it overestimated itself... ☐
- d. ..he (the young tree) seems to have overestimated his abilities.. ☐
- e. ...but they (the young trees) look like they don't know their ability... ☐
- f. It looked like the grass and the trees overestimated their ability.. ☐
- g. It looks like there is a big energy.. ☐

Comments/introspection on translation of Phrase 3.D:

Text 3, Phrase E

还要荫庇和哺育那按照大自然的安排将要取代它的松林。

- a. ..and it (the young tree) has to protect and feed the forest which will replace it, according to the arrangements of nature. ☐
- b. ..and it (the young tree) has to protect and feed the forest it is going to replace, according to the arrangements of nature. ☐
- c. ..and it (the young tree) has to protect and feed the forest which is going to be replaced by nature. ☐
- d. ..and it (the young tree) has to protect the arrangements of nature by replacing the forest. ☐
- e. ..and it (the young tree) has to protect and feed the forest according to the arrangements of nature. ☐
- f. The young tree has to protect and feed nature and nature will arrange to replace it. ☐
- g. The young tree is protected by nature and produces the forest. ☐
- h. The young tree is protected and fed by the forest, by arrangement with the natural surroundings. ☐
- I. We need to follow the arrangements of nature to take care of the forest. ☐
- j. ...it depends on the way the young trees have grown up. ☐

Comments /introspection on translation of Phrase 3.E:

**Appendix III: examination papers**

# London Examinations GCE

Advanced Level

Tuesday 19 May 1998 - Afternoon

Chinese

Paper 1

Time: 3 hours

## Instructions to Candidates

In the boxes on the answer book, write the name of the examining body (London Examinations), your centre number, candidate number, the subject/module title, the paper reference, your surname, other names and signature.

The paper reference is shown in the top left hand corner. You should write the one for which you have been entered.

Answer ALL questions in section A, and ANY THREE questions in Section B.

Each Chinese passage for translation into English is printed twice in this question paper, once in full characters ( 繁體字 ), and once in simplified characters ( 简体字 ).

**YOU MUST TRANSLATE EACH PASSAGE ONLY ONCE.**

## Information for Candidates

You are reminded of the necessity for good English and orderly presentation in your answers.

H0598 T9609

## Section A. Translation

Answer ALL THREE questions in this section.

1. Translate into English:

*Chinese version in full characters:*

一些摹擬魯迅先生雜文的人，往往忘卻了這些雜文產生的時間和空間，目的和對象。在他的雜文中，有些地方不得已的“咬文嚼字”，“拐彎抹角”，“引經據典”，“聲東擊西”，“文白混用”，或故意“含糊其辭”，都是當時的年代、環境的產物，連先生自己也認為詬病的。如今我們不能盲目地效仿，應該只是在需要的時候才能夠用一用它。我們所主要該學的，卻是那及時磨練武器、精通武藝、以及臨敵無畏的精神。

*Chinese version in simplified characters:*

一些摹擬魯迅先生雜文的人，往往忘卻了這些雜文產生的時間和空間，目的和對象。在他的雜文中，有些地方不得已的“咬文嚼字”，“拐彎抹角”，“引經據典”，“聲東擊西”，“文白混用”，或故意“含糊其辭”，都是當時的年代、環境的產物，連先生自己也認為詬病的。如今我們不能盲目地效仿，應該只是在需要的時候才能夠用一用它。我們所主要該學的，卻是那及時磨練武器、精通武藝、以及臨敵無畏的精神。

(20 marks)

- 
2. Translate into English:

*Chinese version in full characters:*

我選了一粒瓜子，送交臼齒去咬。不幸這瓜子太燥，我用力又太猛，“格”的一響，玉石不分，咬成無數的碎塊，我只得把黏著唾液的碎塊吐在手心里，用心挑選，剔出殼的碎塊，然後用舌尖舐食瓜仁的碎塊。然而這挑選頗不容易，因為殼的碎塊的一面也是白色的，與瓜仁無異，我誤認為全是瓜仁而舐進口中去嚼，其味雖非嚼蠟，卻等於嚼砂。殼的碎片緊緊地嵌進牙縫裏，找不到牙簽就無法取出。碰到這種釘子的時候，我就下決心，從此戒絕瓜子。

*Chinese version in simplified characters:*

我选了一粒瓜子，送交臼齿去咬。不幸这瓜子太燥，我用力又太猛，“格”的一响，玉石不分，咬成无数的碎块，我只得把黏着唾液的碎块吐在手心里，用心挑选，剔出壳的碎块，然后用舌尖舐食瓜仁的碎块。然而这挑选颇不容易，因为壳的碎块的一面也是白色的，与瓜仁无异，我误认为全是瓜仁而舐进口中去嚼，其味虽非嚼蜡，却等于嚼砂。壳的碎片紧紧地嵌进牙缝里，找不到牙签就无法取出。碰到这种钉子的时候，我就下决心，从此戒绝瓜子。

(20 marks)

3. Translate into English:

*Chinese version in full characters:*

當幼樹從松林的廢墟上破土而出，在天邊舞動那皎潔的苗條，這不像是又回到了破曉的黎明，看到了新一代的大森林露出了它燦爛奪目的曙光嗎！憑著它那纖弱蒼白的一束嫩條，幼樹是多麼堅毅和勇敢的先驅者，似乎是不自量力竟要在一片荒漠貧瘠的土地上從頭拓荒，要重建一座比已經有過的都更加偉大的森林。它又是怎樣達觀和睿智的開拓者，還要蔭庇和哺育那按照大自然的安排將要取代它的松林。

*Chinese version in simplified characters:*

当幼树从松林的废墟上破土而出，在天边舞动那皎洁的苗条，这不象是又回到了破晓的黎明，看到了新一代的大森林露出了它灿烂夺目的曙光吗！凭着它那纤弱苍白的一束嫩条，幼树是多么坚毅和勇敢的先驱者，似乎是不自量力竟要在一片荒漠贫瘠的土地上从头拓荒，要重建一座比已经有过的都更加伟大的森林。它又是怎样达观和睿智的开拓者，还要荫庇和哺育那按照大自然的安排将要取代它的松林。

(15 marks)

*Turn over*

9609/01

# Edexcel GCE

## Chinese

Paper 1

Advanced Level

Wednesday 24 May 2000 – Afternoon

Time: 3 hours

**N4684**

**Materials required for examination:**

Answer Book (AB02)  
Additional Answer Sheet (AS01).

**Items included with these question papers:**

Nil.

### Instructions to Candidates

In the boxes on the Answer Book, write the name of the Examining Body (Edexcel), your Centre Number, Candidate Number, the Subject/Module title, the Paper reference, your Surname, Other names and Signature.

The Paper reference is shown in the top left-hand corner.

Answer ALL questions in Section A, and ANY THREE questions in Section B.

Each Chinese passage for translation into English is printed twice in this question paper, once in full characters (繁體字), and once in simplified characters (简体字).

**YOU MUST TRANSLATE EACH PASSAGE ONLY ONCE.**

Answer your questions in the Answer Book. Additional Answer Sheets may be used.

### Information for Candidates

The marks for the various parts of questions are shown in round brackets: e.g. (20 marks).

All blank pages are indicated.

### Advice to Candidates

You are reminded of the necessity for good English and orderly presentation in your answers.

*Turn over*



## Section A. Translation

Answer ALL THREE questions in this section.

### 1. Translate into English:

*Chinese version in full characters:*

帶著中學生的挑戰，我們去請教了教育家和管教育的行政長官，把孩子們七嘴八舌的議論，一古腦端給了他們。得到的回答是：教育內容陳舊，教育方法死板，不受歡迎。

一位副處長說：“現在的教育格局是千軍萬馬過獨木橋，無視一百個同齡中學生只能有十個人升高中、三個人上大學這樣的事實，片面追求升學率。高考指揮棒不僅指揮著只佔全國中學百分之三的重點中學，也指揮著那些被稱為第三世界的一般中學。一些升學率低的學校名聲掃地，社會冷眼相看，家長持歧視態度，學校如何辦得下去呢？”

國家教委發言人言談中也充滿焦慮：“專家是人才，有一定專長的中學技術人員也是人才，經濟的發展需要多層次多門類的人才，需要提高全民族的文化素質。倘若光培養尖子，民族振興何日可待？”

*Chinese version in simplified characters:*

带着中学生的挑战，我们去请教了教育家和管教育的行政长官，把孩子们七嘴八舌的议论，一古脑端给了他们。得到的回答是：教育内容陈旧，教育方法死板，不受欢迎。

一位副处长说：“现在的教育格局是千军万马过独木桥，无视一百个同龄中学生只能有十个人升高中、三個人上大学这样的事实，片面追求升学率。高考指挥棒不仅指挥着只占全国中学百分之三的重点中学，也指挥着那些被称为第三世界的一般中学。一些升学率低的学校名声扫地，社会冷眼相看，家长持歧视态度，学校如何办得下去呢？”

国家教委发言人言谈中也充满焦虑：“专家是人才，有一定专长的中学技术人员也是人才，经济的发展需要多层次多门类的人才，需要提高全民族的文化素质。倘若光培养尖子，民族振兴何日可待？”

(20 marks)

2. Translate into English:

*Chinese version in full characters:*

· · · · · 小蔣在大陸好歹是個文化圈裏熏陶出來的，當天晚上酒樓裏燈火輝煌，因為過年的緣故，前來歡宴的顧客如雲。小蔣器宇軒昂地率領一溜兒高矮不齊的隊伍踏上樓梯，揮手叫來服務員，選中一張圓桌圍席入座。這是一次奇特的晚宴，坐在主人位置上的是珠環翠繞的小蔣，身旁偎著她的那個穿著珠繡綢緞中式絲棉襖的小少爺，而被宴請的主賓則是一位面色淒苦的中年婦女，旁邊依次排列著四個蔫頭搭腦的男孩和女孩。小蔣央求她那胡攪蠻纏的寶貝賜給片刻的寧靜，忙著替另外的四個孩子夾菜盛湯，而她自己光顧滔滔不絕地說話，顧不上吃飯。

*Chinese version in simplified characters:*

· · · · · 小蔣在大陸好歹是个文化圈里熏陶出来的，当天晚上酒楼里灯火辉煌，因为过年的缘故，前来欢宴的顾客如云。小蔣器宇轩昂地率领一溜儿高矮不齐的队伍踏上楼梯，挥手叫来服务员，选中一张圆桌围席入座。这是一次奇特的晚宴，坐在主人位置上的是珠环翠绕的小蔣，身旁偎着她的那个穿着珠绣绸缎中式丝棉袄的小少爷，而被宴请的主宾则是一位面色凄苦的中年妇女，旁边依次排列着四个蔫头搭脑的男孩和女孩。小蔣央求她那胡搅蛮缠的宝贝赐给片刻的宁静，忙着替另外的四个孩子夹菜盛汤，而她自己光顾滔滔不绝地说话，顾不上吃饭。

(20 marks)

---